The Application of Dynamic Assessment in College English Writing Instruction: A Multi-Dimensional Analysis of Students' Perceptions and Proficiency-Based Differences

Feng Jinghan[1,2*], Noor Mala Ibrahim[1], Shao Wenjing[3], Zhao Lijin[1], Ji Xiaomeng[3]

[1]School of Education, University Technology Malaysia, Johor Bahru, Malaysia

[2]Department of Foreign Language Education and Teaching, Hebei Finance University, Hebei Province, China

[3]Language Academy, University Technology Malaysia, Johor Bahru, Malaysia

***Corresponding Author**

DOI: https://dx.doi.org/10.47772/IJRISS.2025.91100159

Received: 16 November 2025; Accepted: 22 November 2025; Published: 03 December 2025

ABSTRACT

College English writing instruction in China has long been plagued by challenges such as the ineffectiveness of traditional static assessment, insufficient personalized feedback due to the disproportionate teacher-student ratio, and students' difficulties in improving writing performance and behaviors (Wei, 2010; Ning, 2021; Zhang, 2023). Dynamic Assessment (DA), which integrates instruction and assessment through contingent mediation based on Vygotsky's Sociocultural Theory and Zone of Proximal Development, has been proposed as a potential solution to address these issues (Lantolf & Poehner, 2004). This study aims to explore non-English major students' perceptions of DA in college English writing instruction and examine whether there are significant differences in these perceptions across students of high, medium, and low proficiency levels. A mixed-method research design was adopted, with 196 second-year undergraduate students from Hebei Finance University completing a modified Tripod Student Perception Survey covering seven dimensions: CARE, CONTROL, CLARIFY, CHALLENGE, CAPTIVATE, CONFER, and CONSOLIDATE. Results indicate that students hold overall positive perceptions of DA, with CONSOLIDATE (M=4.5) and CARE (M=4.48) receiving the highest scores, while CONTROL (M=3.91) scored the lowest. Kruskal-Wallis H tests revealed significant differences only in the CONFER dimension (p=0.016), with post-hoc Mann-Whitney U tests showing that high-proficiency students scored significantly higher than medium-proficiency students (p=0.004). These findings provide empirical support for the applicability of DA in group classroom settings and offer insights for optimizing college English writing instruction tailored to students of different proficiency levels.

Keywords: Dynamic Assessment, English writing instruction, students' perceptions

INTRODUCTION

Since the implementation of the Reform and Opening-up Policy in 1978, English education has occupied a crucial position in China's education system, serving as a compulsory course in primary, secondary, and higher education institutions (Cai, 2005; Yu & Yuan, 2005). This emphasis stems not only from English's status as an international language but also from China's commitment to integrating into the global market and cultivating talents capable of meeting the demands of globalization (Wang, 2016). In line with the revised College English Teaching Guidelines (2020) issued by the Ministry of Education, higher education institutions are under continuous pressure to enhance the quality of college English teaching, particularly given the significance of the College English Test Band-4 (CET-4) for students' academic and career development (He, 2020; Jin et al., 2022).

Among the five language skills (listening, speaking, reading, writing, and translating), writing has long been recognized as one of the most challenging for Chinese learners (Xu, 2023). Despite substantial investments of

time and effort by teachers and students, outcomes remain unsatisfactory, with students facing difficulties in content development, structural organization, language authenticity, and revision strategies (Zhang, 2023). These issues are exacerbated in non-English major classes, where university expansion has led to a shortage of English teachers, resulting in limited opportunities for personalized feedback and guidance (Liu & Qi, 2024; Huang & Long, 2023).

A key contributing factor to these challenges is the persistence of traditional static assessment approaches in college English writing instruction. This model separates instruction from assessment, focusing solely on evaluating final products with minimal or no feedback, leaving students unmotivated and unsure of how to revise their work (Peng, 2022; Yang, 2024). In contrast, Dynamic Assessment (DA) offers a promising alternative by integrating instruction and assessment through mediated interaction, aiming to promote learner development within their Zone of Proximal Development (ZPD) (Lantolf & Poehner, 2004). Rooted in Vygotsky's Social Cultural Theory, DA emphasizes the role of social interaction and scaffolded support in fostering cognitive development, making it particularly relevant to writing—a complex, iterative process involving planning, drafting, and revision (Khanlarzadeh & Nemati, 2016).

While existing research has demonstrated DA's effectiveness in individualized settings (Aljaafreh & Lantolf, 1994; Shrestha & Coffin, 2012), its application in group classroom contexts remains underexplored. Furthermore, few studies have examined students' perceptions of DA, especially across different proficiency levels, which is critical for understanding its practical acceptance and optimizing implementation.

This study therefore addresses two core research questions:

1. What are non-English major students' perceptions of DA in college English writing instruction, as reflected in the seven dimensions of the Tripod framework (CARE, CONTROL, CLARIFY, CHALLENGE, CAPTIVATE, CONFER, CONSOLIDATE)?
2. Are there significant differences in these perceptions between students of high, medium, and low English proficiency levels?

By exploring these questions, this research aims to contribute to the empirical literature on DA in group writing classrooms, provide insights into students' acceptance of DA, and offer practical recommendations for enhancing college English writing instruction.

LITERATURE REVIEW

This part dates back to the literature of writing instruction and assessment, the theoretical foundations of DA, DA in L2 writing instruction and students' perception of it.

Research on Writing Instruction and Assessment

Writing instruction has long been a focal point of second language acquisition research, with two dominant paradigms shaping pedagogical practices: the product approach and the process approach. The product approach, rooted in traditional rhetoric, emphasizes the final written product, focusing on grammatical accuracy, lexical appropriateness, and textual structure. Teachers employing this approach typically provide model texts for imitation, with assessment centered on summative evaluation of the finished work. However, this model has been criticized for neglecting the cognitive and metacognitive processes involved in writing, such as idea generation, revision, and reflection, leading to limited improvement in students' long-term writing competence (Ning, 2021).

In contrast, the process approach views writing as a dynamic, iterative activity involving pre-writing, drafting, revising, and editing. This approach emphasizes the development of writing strategies and self-regulation, with assessment integrated into the learning process to provide timely feedback. Despite its advantages, process-oriented instruction in Chinese college English classrooms faces challenges, including large class sizes, limited teacher-student interaction, and the persistence of static assessment methods that prioritize scores over growth (Wei, 2010).

Theoretical Foundations and Core Concepts of Dynamic Assessment (DA)

Dynamic Assessment (DA) emerged as a response to the limitations of static assessment, drawing on Vygotsky's Social Cultural Theory (SCT) and the concept of the Zone of Proximal Development (ZPD). SCT posits that cognitive development is socially mediated, with learning occurring through interaction with more knowledgeable others (Vygotsky, 1978). The ZPD refers to the gap between a learner's current independent ability and their potential development with guidance, forming the theoretical basis for DA's integration of instruction and assessment.

DA is defined as "a procedure that integrates assessment and instruction into a seamless, unified activity aimed at promoting learner growth through appropriate forms of mediation" (Lantolf & Poehner, 2004). Its core characteristics include: (1) the fusion of teaching and assessment; (2) contingent mediation tailored to learners' ZPD; and (3) a focus on developmental potential rather than fixed ability. Two primary approaches to DA are identified: interventionist DA, which uses pre-scripted mediation, and interactionist DA, which employs flexible, dialogue-driven support.

DA in L2 Writing Instruction: Progress and Gaps

Research on DA in second language (L2) writing has demonstrated its potential to enhance learners' performance. Early studies, such as Aljaafreh and Lantolf (1994), showed that contingent mediation improved learners' grammatical accuracy and self-correction abilities in writing. More recent studies, including Shrestha and Coffin (2012), highlighted DA's role in promoting academic writing skills and conceptual knowledge through iterative feedback.

However, significant gaps remain in the literature. First, most DA studies in L2 writing focus on individualized settings, with limited exploration of its application in group classroom contexts (Poehner, 2009). Given that college English classes in China typically consist of 40-50 students, this gap hinders the practical application of DA. Second, existing research primarily examines DA's impact on writing products (e.g., scores, error correction) rather than learners' perceptions of the approach, which are critical for sustained engagement. Third, few studies investigate differences in DA perceptions across learners of varying proficiency levels, despite evidence that high, medium, and low-level students may respond differently to mediation.

Students' Perceptions of Assessment in L2 Writing

Learner perceptions shape their engagement with instructional practices, yet research on how students perceive DA in writing is scarce. Studies on traditional assessment note that summative feedback often leaves students feeling passive and uncertain about revision strategies (Yang, 2024). In contrast, process-oriented assessment, such as DA, has been associated with increased motivation, as learners perceive feedback as supportive rather than judgmental.

The Tripod framework (Ferguson, 2010), which includes seven dimensions of effective teaching (CARE, CONTROL, CLARIFY, CHALLENGE, CAPTIVATE, CONFER, CONSOLIDATE), provides a useful lens to explore students' perceptions of DA. This framework aligns with DA's emphasis on interaction, mediation, and growth, making it suitable for analyzing how learners experience DA in writing classrooms.

In summary, while DA shows promise for addressing limitations in college English writing instruction, research on its application in group settings, learners' perceptions, and proficiency-based differences remains underdeveloped. This study addresses these gaps by examining non-English majors' perceptions of DA across the seven Tripod dimensions and exploring variations among high, medium, and low-proficiency students.

METHODOLOGY

This part introduces the research design, participants, research instruments and data analysis process.

Research Design

This study employed a mixed-methods design, with a primary focus on quantitative data to examine students' perceptions of Dynamic Assessment (DA) in college English writing instruction, supplemented by qualitative

data from open-ended questionnaire comments to elaborate on the quantitative results. As emphasized by Creswell (2003), mixed-methods research combines the strengths of quantitative and qualitative approaches, offering a more comprehensive understanding of research problems than either method alone. In this study, quantitative data from the questionnaire provided an overview of students' overall perceptions and group differences, while qualitative data from open-ended responses enriched the analysis by explaining the reasons behind specific perceptions (e.g., why certain dimensions of DA received higher ratings).

Participants

The participants were 196 second-year undergraduate students majoring in accounting at Hebei Finance University, aged 18–20, with 12 years of English learning experience. They were selected using simple random sampling from a total population of 400 students. All participants were non-English majors, native Chinese speakers, and had completed high school education through the National College Entrance Examination ("Gaokao").

To explore potential differences in perceptions across proficiency levels, participants were divided into three groups based on their pre-test scores using Jacobs et al.'s (1981) Scale for Assessing Second Language Writing Proficiency:

1. High proficiency: Scores ≥ 90 (n=60)
2. Medium proficiency: Scores 67–89 (n=67)
3. Low proficiency: Scores < 67 (n=66)

This grouping aligns with the College English Teaching Guidelines (2020), which categorize college students' English proficiency into three levels, ensuring the sample's relevance to real-world teaching contexts.

Research Instrument

The primary research instrument was a modified version of the Tripod Student Perception Survey (Ferguson, 2010), which was adapted to focus on students' perceptions of DA in writing instruction. The questionnaire consisted of 36 items across seven dimensions, designed to measure key aspects of DA implementation:

1. CARE: Perceptions of teacher-student emotional connection and support.
2. CONTROL: Perceptions of classroom management and order.
3. CLARIFY: Perceptions of teachers' ability to explain writing concepts clearly.
4. CHALLENGE: Perceptions of appropriate academic pressure and motivation.
5. CAPTIVATE: Perceptions of the engaging nature of DA activities.
6. CONFER: Perceptions of opportunities to express opinions and ideas.
7. CONSOLIDATE: Perceptions of knowledge integration and skill reinforcement.

Responses were recorded on a 5-point Likert scale (1="Totally Untrue" to 5="Totally True"). Five negative items (Items 5, 6, 7, 13, 24) were reverse-coded to ensure consistent scoring. The questionnaire was adjusted to reflect writing-specific contexts (e.g., "My teacher seems to know if something is bothering me while I'm writing" instead of the original "My teacher seems to know if something is bothering me") to enhance relevance.

Additionally, an open-ended question at the end of the questionnaire—"What aspect of the questionnaire resonates with you the most?"—collected qualitative data to supplement the quantitative results. This allowed students to elaborate on their experiences, providing context for high or low ratings in specific dimensions (e.g., explaining why CONSOLIDATE received the highest mean score).

Data Analysis

Quantitative data from the questionnaire were analyzed using the following statistical procedures.

Descriptive statistics: Mean scores and standard deviations were calculated for each of the seven dimensions to identify overall trends in students' perceptions. This included examining central tendency (mean, median,

mode) and dispersion (standard deviation) to characterize the distribution of responses.

Normality testing: The Kolmogorov-Smirnov test was used to check if data conformed to a normal distribution. This determined the suitability of parametric or non-parametric tests for further analysis.

Non-parametric tests: Given potential non-normal distribution, the Kruskal-Wallis H test was employed to examine differences in perceptions across the three proficiency groups. A post-hoc Mann-Whitney U test with Bonferroni correction (α=0.017) was used to identify specific pairwise differences between groups.

Qualitative data from open-ended questionnaire comments were analyzed using thematic analysis, following Braun and Clarke's (2006) six-step framework: Familiarization: Transcribing and repeatedly reading the comments to gain a holistic understanding of the data.

Generating initial codes: Systematically labeling segments of text to capture key ideas (e.g., "teacher summaries," "classroom distractions").

Searching for themes: Grouping codes into broader themes that reflect shared patterns (e.g., "perceived effectiveness of knowledge consolidation").

Reviewing themes: Validating that themes are consistent with the raw data and refining their scope.

Defining and naming themes: Clarifying the core meaning of each theme and assigning

The analysis focused on identifying recurring patterns in students' responses to explain the quantitative results. Qualitative analysis can explore the reasons behind high or low scores in a certain dimension, as well as help understand why students at different proficiency levels hold varying views on a specific dimension.

RESULTS AND ANALYSIS

A total of 196 students from Hebei Finance University participated in the study, with 193 valid questionnaires returned (response rate = 98.5%).

Overall Perceptions Across Seven Dimensions

Descriptive statistics (mean and standard deviation) for the seven dimensions are presented in Table 1, revealing students' overall positive attitudes toward DA, with variations in the degree of approval.

Table 1. Mean scores and standard deviations of the seven dimensions

Dimension	Mean	SD	Rank
CONSOLIDATE	4.50	0.65	1
CARE	4.48	0.63	2
CHALLENGE	4.42	0.71	3
CLARIFY	4.32	0.74	4
CONFER	4.12	0.96	5
CAPTIVATE	4.03	0.93	6
CONTROL	3.91	1.02	7

Consolidate: Perceptions of Knowledge Integration and Skill Reinforcement

As the highest-scoring dimension (M=4.50, SD=0.65), CONSOLIDATE reflects students' perceptions of how DA facilitates the integration of writing knowledge (e.g., structural logic, error correction) and reinforces skills for long-term use (Ferguson, 2010). This dimension aligns with DA's core goal—moving beyond "one-time feedback" to help learners internalize strategies (Lantolf & Poehner, 2004)—and addresses a key limitation of traditional writing assessment, where students often fail to apply feedback to future tasks (Yang, 2024).

Table 2 presents descriptive statistics for CONSOLIDATE items, with all receiving high mean scores (4.34–4.67) and over 85% of students selecting "Strongly Agree" or "Agree."

Table 2. Descriptive Statistics for CONSOLIDATE Items

Item Number	Survey Item	SA (5) n (%)	A (4) n (%)	NS (3) n (%)	D (2) n (%)	SD (1) n (%)	Mean	SD
33	My teacher takes the time to summarize what we learn in each writing class.	133 (68.9)	56 (29.0)	4 (2.1)	0 (0)	0 (0)	4.67	0.51
36	The revisions that my teacher and classmates make to other students' essays help me understand how my essay should be improved.	117 (60.6)	67 (34.7)	5 (2.6)	4 (2.1)	0 (0)	4.54	0.65
35	My review of other students' essays in my writing class shows me what is wrong with my own essays.	106 (54.9)	75 (38.9)	8 (4.1)	4 (2.1)	0 (0)	4.47	0.68
34	My teacher checks to make sure we understand what she is teaching us in the writing class.	90 (46.6)	84 (43.5)	14 (7.3)	5 (2.6)	0 (0)	4.34	0.72

Note: SA=Strongly Agree; A=Agree; NS=Not Sure; D=Disagree; SD=Strongly Disagree. Total valid responses=193.

Item 33 scored highest (M=4.67, SD=0.51), with 97.9% of students agreeing that in-class summaries (e.g., recapitulation of common errors, key writing structures) helped them consolidate knowledge. This reflects DA's focus on "process-oriented review"—unlike traditional classes where summaries are rare, DA uses post-task recapitulations to reinforce what students learned (Shabani, 2018).

Item 36 (M=4.54, SD=0.65) and Item 35 (M=4.47, SD=0.68) highlighted the value of peer/teacher revision of sample essays: 95.3% and 93.8% of students agreed this activity helped them identify their own mistakes. For example, analyzing a classmate's disorganized paragraph taught students to check their own essay structure—aligning with DA's goal of fostering self-regulation (Aljaafreh & Lantolf, 1994).

Item 34 had the lowest mean (M=4.34, SD=0.72), with 7.3% of students selecting "Not Sure." This may be due to occasional gaps in teacher check-ins (e.g., moving to the next task before all students grasp a concept), particularly in large classes (Liu & Qi, 2024).

Thematic analysis of open-ended responses (n=193) confirmed students' positive perceptions, identifying two core themes:

In-class summaries clarify key points: 72% of students emphasized that teacher summaries (e.g., listing common errors like "subject-verb inconsistency") helped them retain knowledge. A high-proficiency student noted, "The teacher's end-of-class summary lets me check what I missed—now I actively avoid mistakes I used to make."

Peer revision builds self-correction skills: 65% of students reported that analyzing classmates' essays taught them to "spot errors in my own work." A low-proficiency student commented, "I used to rely on the teacher to correct my essays, but now I can find simple mistakes (like wrong prepositions) by reviewing others' work—it's like practicing for my own writing."

Kruskal-Wallis H tests showed no significant differences in CONSOLIDATE perceptions across high (M=4.55, SD=0.61), medium (M=4.48, SD=0.67), and low-proficiency (M=4.46, SD=0.68) groups (χ^2=1.312, df=2, p=0.519 > 0.05). This uniformity stems from DA's adaptive mediation: high-proficiency students consolidated advanced skills (e.g., logical argumentation), while low-proficiency students focused on basic skills (e.g., grammar error correction)—both groups benefited from knowledge integration.

The CONSOLIDATE dimension demonstrates DA's strength in helping students internalize writing knowledge and skills. Students particularly value in-class summaries and peer revision, which address the "feedback disconnection" of traditional assessment. The only improvement area is enhancing teacher check-ins to ensure all students grasp key concepts—though this does not diminish the dimension's overall effectiveness. For college English writing instruction, CONSOLIDATE confirms DA's ability to support long-term skill growth, not just short-term task completion.

Care: Perceptions of Teacher-Student Emotional Connection and Support

As the second-highest scoring dimension (M=4.48, SD=0.63), CARE captures students' perceptions of emotional rapport, individualized attention, and supportive interactions with teachers during DA-based writing instruction (Ferguson, 2010). Aligned with Vygotsky's Social Cultural Theory— which emphasizes learning through supportive social contexts (Lantolf & Thorne, 2006)—this dimension reflects DA's strength in fostering trust and emotional safety, critical for reducing writing anxiety among non-English majors (Zhang, 2023).

Table 3 presents descriptive statistics for CARE items, with all receiving high mean scores (4.28–4.75) and strong student approval.

Table 3. Descriptive Statistics for CARE Items

Item Number	Survey Item	SA (5) n (%)	A (4) n (%)	NS (3) n (%)	D (2) n (%)	SD (1) n (%)	Mean	SD
1	My teacher in DA writing instruction makes me feel that she really cares about me.	90 (46.6)	94 (48.7)	8 (4.2)	1 (0.5)	0 (0)	4.41	0.60
2	My teacher seems to know if something is bothering me while I'm writing and revising my work.	87 (45.1)	74 (38.3)	32 (16.6)	0 (0)	0 (0)	4.28	0.73
3	My teacher really tries to understand how students feel about writing.	146 (75.7)	45 (23.3)	2 (1.0)	0 (0)	0 (0)	4.75	0.46

Item 3 scored highest (M=4.75, SD=0.46), with 99% of students agreeing teachers prioritized understanding their feelings about writing. This reflects DA's dialogue-driven mediation—unlike traditional static assessment, DA requires ongoing check-ins (e.g., asking about revision challenges) that signal attentiveness to emotional experiences (Poehner, 2008).

Item 1 (M=4.41, SD=0.60) showed 95.3% of students felt teachers conveyed care, primarily through individualized actions (e.g., circulating during writing tasks, one-on-one guidance). Only 0.5% disagreed, confirming DA's interactive structure fosters emotional connection (Wallace & Chhuon, 2014).

Item 2 had the lowest mean (M=4.28, SD=0.73), with 16.6% of students selecting "Not Sure." This likely stems from large class sizes (Liu & Qi, 2024), which limit teachers' ability to detect every student's discomfort, especially among low-proficiency learners hesitant to express struggles.

Thematic analysis of open-ended responses (n=193) corroborated quantitative results, identifying two core themes:

Individualized mediation as a marker of care: 68% of students noted DA's personalized support (e.g., targeted hints, one-on-one outline help) made them feel valued. A low-proficiency student commented, "The teacher noticed I struggled with complex sentences and gave simple examples—she cared about my problem, not just the class."

Emotional safety from non-judgmental feedback: 32% of students highlighted DA's growth-focused approach (e.g., implicit hints before correction) reduced anxiety. A medium-proficiency student stated, "I'm not scared to admit mistakes now—the teacher cares more about learning than criticizing."

Kruskal-Wallis H tests showed no significant differences in CARE perceptions across high (M=4.52, SD=0.58), medium (M=4.47, SD=0.65), and low-proficiency (M=4.45, SD=0.66) groups (χ^2=0.380, df=2, p=0.827 > 0.05). This aligns with DA's design—its integrated assessment-instruction ensures all students receive attention, with high-proficiency learners valuing idea-focused care and low-proficiency learners appreciating reduced anxiety (Ning, 2021).

The CARE dimension confirms DA's effectiveness in fostering emotional rapport. Students strongly recognize teachers' efforts to understand their feelings and convey care, with minimal proficiency-based differences. The only improvement area is enhancing teachers' ability to detect individual discomfort in large classes (e.g., regular check-ins), which would further strengthen this dimension.

Challenge: Perceptions of Appropriate Academic Pressure and Motivation

Ranking third among the seven dimensions (M=4.42, SD=0.71), CHALLENGE reflects students' perceptions of whether DA-based writing instruction provides appropriately challenging tasks and sufficient motivation to engage with them (Ferguson, 2010). This dimension aligns with Vygotsky's "Zone of Proximal Development" (ZPD), which emphasizes that optimal learning occurs when tasks are challenging enough to require effort but achievable with mediation—avoiding both frustration (too difficult) and complacency (too easy) (Lantolf & Thorne, 2006). For non-English majors, who often struggle with writing motivation (Zhang, 2023), balanced challenge is critical for sustaining engagement.

Table 4 presents descriptive statistics for CHALLENGE items, with all receiving positive mean scores (4.25–4.58) and over 80% of students expressing agreement.

Table 4. Descriptive Statistics for CHALLENGE Items

Item Number	Survey Item	SA (5) n (%)	A (4) n (%)	NS (3) n (%)	D (2) n (%)	SD (1) n (%)	Mean	SD
18	My writing teacher makes us think hard, but not so hard that we give up.	112 (58.0)	65 (33.7)	13 (6.7)	3 (1.6)	0 (0)	4.50	0.66
19	My teacher gives us writing assignments that make us want to do our best.	108 (55.9)	69 (35.7)	12 (6.2)	4 (2.1)	0 (0)	4.48	0.68
20	We learn a lot in writing class because it's hard enough to keep us interested.	99 (51.3)	74 (38.3)	15 (7.8)	5 (2.6)	0 (0)	4.38	0.72
21	The writing assignments in this class are challenging but feel doable.	95 (48.2)	70 (36.3)	20 (10.4)	8 (4.1)	0 (0)	4.25	0.81

Item 18 scored highest (M=4.50, SD=0.66), with 91.7% of students agreeing that teachers balanced difficulty—tasks required effort but did not lead to discouragement. This reflects DA's core strength: mediation (e.g., hints, scaffolding) adjusts challenge levels in real time, ensuring tasks stay within students' ZPD (Poehner, 2008).

Item 19 (M=4.48, SD=0.68) highlighted strong student motivation, with 91.6% reporting that assignments inspired effort. DA's focus on growth (rather than grades) likely drives this—students are motivated by visible progress (e.g., "I can now write a clear topic sentence, which I couldn't do before") (Shabani, 2018).

Item 21 had the lowest mean (M=4.25, SD=0.81), with 10.4% of students selecting "Not Sure" and 4.1% disagreeing. This suggests a small subset found assignments either too challenging (low-proficiency) or insufficiently stimulating (high-proficiency), highlighting the difficulty of tailoring tasks to diverse skill levels (Liu & Qi, 2024).

Thematic analysis of open-ended responses (n=193) identified two core themes supporting quantitative results:

Balanced difficulty fosters persistence: 63% of students emphasized that tasks were "hard but manageable." A medium-proficiency student noted, "The essay topics are challenging—last week we wrote about environmental

issues, which I didn't know much about. But the teacher gave us research hints, so I didn't give up. Finishing it made me proud."

Growth-focused feedback enhances motivation: 58% of students linked challenge to a sense of progress. A low-proficiency student commented, "I used to hate writing because it was too hard. Now, the teacher starts with simple tasks (like writing a paragraph) and slowly makes them harder. When I see I can do more than before, I want to try even harder."

Kruskal-Wallis H tests revealed no significant differences in CHALLENGE perceptions across high (M=4.45, SD=0.69), medium (M=4.43, SD=0.73), and low-proficiency (M=4.38, SD=0.71) groups (χ^2=0.521, df=2, p=0.771 > 0.05). This uniformity arises because DA's adaptive mediation tailors challenge to each level: high-proficiency students tackle complex tasks (e.g., argumentative essays with counterpoints), while low-proficiency students focus on foundational skills (e.g., sentence structure)—both groups experience appropriate difficulty.

The CHALLENGE dimension confirms DA's effectiveness in balancing academic pressure and motivation. Students strongly agree that tasks are challenging yet achievable, with mediation ensuring alignment with their ZPD. While a small minority found tasks misaligned with their skill level, the overall consensus highlights DA's advantage over traditional instruction, where one-size-fits-all assignments often fail to motivate. For college English writing, CHALLENGE demonstrates that well-calibrated difficulty, paired with supportive mediation, fosters sustained engagement and effort.

Clarify: Perceptions of Teachers' Ability to Explain Writing Concepts Clearly

Ranking fourth among the seven dimensions (M=4.32, SD=0.74), CLARIFY captures students' perceptions of whether teachers can identify their knowledge gaps in writing (e.g., structural confusion, grammatical ambiguity) and provide clear, diversified explanations to address them (Ferguson, 2010). This dimension aligns with DA's core principle of "contingent mediation"—teachers adjust their explanations based on students' real-time responses, rather than delivering one-size-fits-all lectures (Aljaafreh & Lantolf, 1994), which solves a key limitation of traditional writing instruction where unclear explanations often leave students confused (Wei, 2010).

Table 5. Descriptive Statistics for CLARIFY Items

Item Number	Survey Item	SA (5) n (%)	A (4) n (%)	NS (3) n (%)	D (2) n (%)	SD (1) n (%)	Mean	SD
11	If I don't understand a writing concept, my teacher explains it another way.	128 (66.3)	60 (31.1)	5 (2.6)	0 (0)	0 (0)	4.64	0.53
15	My teacher explains difficult writing concepts (e.g., essay structure) clearly.	99 (51.3)	85 (44.0)	4 (2.1)	3 (1.6)	2 (1.0)	4.43	0.71
14	My teacher has several good ways to explain each writing topic we cover.	92 (47.7)	87 (45.1)	13 (6.7)	1 (0.5)	0 (0)	4.40	0.64
13*	My teacher thinks we understand a writing concept even when we don't.	6 (3.1)	3 (1.6)	40 (20.7)	94 (48.7)	50 (25.9)	3.93	0.90

***Note:** Item 13 is a negative statement, reverse-coded for analysis—higher mean indicates lower agreement with "teacher misjudges understanding."

Item 11 scored highest (M=4.64, SD=0.53), with 97.4% of students agreeing teachers adapt explanations for confusion. This reflects DA's flexible mediation—unlike traditional classes where teachers repeat the same explanation, DA allows teachers to switch methods (e.g., from English to Chinese, from lectures to examples) based on students' reactions (Lantolf & Poehner, 2004).

Item 15 (M=4.43, SD=0.71) and Item 14 (M=4.40, SD=0.64) confirmed teachers' ability to clarify difficult concepts: 95.3% and 92.8% of students agreed, respectively. For example, when explaining "logical coherence," teachers used both sample essays (good vs. poor coherence) and graphic organizers—helping students visualize abstract concepts (Shabani, 2018).

Item 13 (reverse-coded, M=3.93, SD=0.90) had the lowest mean: 74.6% of students disagreed that "teachers misjudge understanding," but 20.7% selected "Not Sure." This uncertainty likely stems from large class sizes (Liu & Qi, 2024)—teachers cannot check every student's understanding in real time, especially for quiet low-proficiency students who avoid asking questions.

Thematic analysis of open-ended responses (n=193) identified two core themes supporting quantitative results:

Diversified explanations resolve confusion: 70% of students emphasized teachers' adaptive methods. A low-proficiency student commented, "I couldn't understand 'topic sentences' at first— the teacher explained it in Chinese, then showed a bad example (no clear main idea) and a good one. Now I get it."

Real-time checks prevent misunderstanding: 55% of students noted teachers' use of questions to verify understanding. A medium-proficiency student stated, "After explaining 'paragraph transitions,' the teacher asked us to point out transitions in a sample essay. If someone got it wrong, she explained again—this made sure no one was left behind."

Kruskal-Wallis H tests showed no significant differences in CLARIFY perceptions across high (M=4.35, SD=0.70), medium (M=4.30, SD=0.75), and low-proficiency (M=4.28, SD=0.76) groups (χ^2=2.363, df=2, p=0.307 > 0.05). This uniformity arises because DA's mediation tailors clarity to proficiency levels: high-proficiency students receive explanations for advanced concepts (e.g., academic tone), while low-proficiency students get guidance on basics (e.g., sentence structure)—both groups perceive explanations as clear.

The CLARIFY dimension confirms DA's effectiveness in addressing students' knowledge gaps through clear, adaptive explanations. Students highly approve of teachers' ability to switch methods and clarify difficult concepts, with minimal proficiency-based differences. The only improvement area is reducing "Not Sure" responses in Item 13—teachers could use more frequent, inclusive checks (e.g., peer discussions to reveal confusion) to avoid misjudging understanding. Overall, CLARIFY demonstrates DA's advantage over traditional instruction, where one-way lectures often fail to resolve individual confusion.

Confer: Perceptions of Opportunities to Express Ideas and Opinions

Ranking fifth among the seven dimensions (M=4.12, SD=0.96), CONFER reflects students' perceptions of whether DA-based writing instruction creates space for them to voice their views on writing tasks, revisions, and learning needs—along with whether teachers value and respond to these expressions (Ferguson, 2010). This dimension ties closely to DA's interactive essence: unlike traditional assessment, where communication is one-way (teacher to student, via grades/comments), DA positions students as active participants in the assessment-instruction process (Poehner, 2009), which is critical for fostering ownership of learning among non-English majors who often feel passive in writing classes (Zhang, 2023).

Table 6 presents descriptive statistics for CONFER items, with most receiving positive mean scores (3.18–4.56) and notable variation in student agreement—particularly for items related to autonomy in activity design.

Table 6. Descriptive Statistics for CONFER Items

Item Number	Survey Item	SA (5) n (%)	A (4) n (%)	NS (3) n (%)	D (2) n (%)	SD (1) n (%)	Mean	SD
32	My teacher respects my ideas and suggestions about writing.	111 (59.1)	75 (38.9)	3 (1.6)	1 (0.5)	0 (0)	4.56	0.55
28	My teacher encourages us to share our thoughts about writing tasks (e.g., essay topics).	97 (50.3)	87 (45.1)	5 (2.6)	4 (2.1)	0 (0)	4.44	0.65
31	I feel comfortable speaking up and sharing my ideas about classmates' essay revisions.	91 (47.2)	74 (38.3)	17 (8.8)	11 (5.7)	0 (0)	4.27	0.85
30	My teacher gives me enough time to explain my ideas about writing.	72 (37.3)	86 (44.6)	23 (12.0)	11 (5.7)	1 (0.5)	4.12	0.87

29	Students get to decide how writing activities are done (e.g., revision order).	29 (15.0)	45 (23.3)	58 (30.1)	54 (28.0)	7 (3.6)	3.18	1.11

Item 32 scored highest (M=4.56, SD=0.55), with 98% of students agreeing teachers respected their ideas. This underscores DA's focus on dialogic mediation—when students proposed alternative revision strategies (e.g., "Can we outline before revising?") or questioned task requirements, teachers acknowledged and integrated valid suggestions, rather than dismissing them (Aljaafreh & Lantolf, 1994).

Item 28 (M=4.44, SD=0.65) and Item 31 (M=4.27, SD=0.85) confirmed strong opportunities for idea-sharing: 95.4% of students felt encouraged to voice thoughts on tasks, and 85.5% felt comfortable sharing revision insights. DA's group-based activities (e.g., collaborative analysis of sample essays) created structured spaces for expression—unlike traditional classes where students rarely speak up (Wei, 2010).

Item 29 had the lowest mean (M=3.18, SD=1.11), with only 38.3% of students agreeing they influenced activity design; 31.6% disagreed, and 30.1% were unsure. This reflects the practical constraint of large classes (Liu & Qi, 2024)—teachers cannot adapt activities to every student's preference, leading to a sense of limited autonomy.

Item 30 (M=4.12, SD=0.87) showed 81.9% of students felt they had time to explain ideas, but 17.7% (12% unsure + 5.7% disagree) felt rushed. This likely stems from time pressure in 45-minute classes, where teachers sometimes cut explanations short to complete planned tasks.

Thematic analysis of open-ended responses (n=193) uncovered three core themes that contextualize quantitative results:

Respect for ideas boosts confidence: 68% of students highlighted that teacher acknowledgment made them more willing to participate. A low-proficiency student noted, "I once suggested using Chinese to brainstorm before writing— the teacher said it was a good idea and let us try. Now I'm not scared to share what I think, even if my English isn't good."

Collaborative revision fosters expression: 56% of students linked group activities to comfort in speaking up. A medium-proficiency student commented, "When we revise a classmate's essay together, I don't feel alone—everyone shares small ideas, and the teacher listens to all of us. It's easier to speak up than when the teacher calls on you one-on-one."

Limited autonomy in activity design causes frustration: 32% of students mentioned wanting more say in how tasks are structured. A high-proficiency student stated, "I wanted to spend more time on argument development, but the teacher stuck to the planned schedule. It feels like we can share ideas, but not change how we learn."

Kruskal-Wallis H tests revealed a significant difference in CONFER perceptions across proficiency groups (χ^2=8.272, df=2, p=0.016 < 0.05)—the only dimension with such variation. Post-hoc Mann-Whitney U tests (Bonferroni-corrected α=0.017) clarified:

High vs. Medium: Significant difference (p=0.004 < 0.017), with high-proficiency students (M=4.35, SD=0.88) scoring higher than medium-proficiency peers (M=4.02, SD=0.95). High-proficiency students were more likely to have their advanced ideas (e.g., "adding counterarguments") adopted, while medium-proficiency students reported more instances of "ideas being overlooked."

High vs. Low: No significant difference (p=0.194 > 0.017). Low-proficiency students (M=4.10, SD=0.92) valued opportunities to share basic needs (e.g., "slower explanation"), which teachers often accommodated—closing the perception gap with high-proficiency learners.

Medium vs. Low: No significant difference (p=0.117 > 0.017). Both groups felt similarly constrained by limited activity autonomy, offsetting differences in idea adoption.

The CONFER dimension highlights DA's strength in creating spaces for students to express ideas, particularly

through teacher respect and collaborative activities. Students across proficiency levels value this opportunity, though high-proficiency learners perceive greater idea adoption than medium-proficiency peers. The key areas for improvement are increasing student influence over activity design (e.g., letting students vote on revision focus) and ensuring sufficient time for explanations—adjustments that would address the lower scores in Items 29 and 30. Overall, CONFER confirms that DA's dialogic nature empowers students to take ownership of their writing learning, even amid the practical limits of large classes.

Captivate: Perceptions of the Engaging Nature of DA Writing Activities

Ranking sixth among the seven dimensions (M=4.03, SD=0.93), CAPTIVATE captures students' perceptions of whether DA-based writing instruction is interesting, engaging, and able to hold their attention—moving beyond the "routine and boring" experience of traditional writing classes (Zhang, 2023). This dimension ties to DA's process-oriented and interactive design: unlike static assessment that centers on isolated essay-writing tasks, DA integrates collaborative activities, real-time feedback, and varied mediation, all of which aim to make writing learning more dynamic (Poehner, 2009). For non-English majors who often report low interest in writing (Ning, 2021), this engagement is key to sustaining long-term participation.

Table 7 presents descriptive statistics for CAPTIVATE items, including a reverse-coded negative item (Item24). Most items reflect positive perceptions, though with more variation than higher-ranked dimensions (e.g., CARE, CONSOLIDATE).

Table 7. Descriptive Statistics for CAPTIVATE Items

Item Number	Survey Item	SA (5) n (%)	A (4) n (%)	NS (3) n (%)	D (2) n (%)	SD (1) n (%)	Mean	SD
26	My teacher designs interesting activities for writing class (e.g., collaborative revision).	86 (44.6)	86 (44.6)	14 (7.3)	7 (3.6)	0 (0)	4.30	0.76
25	I find learning writing through DA enjoyable.	66 (34.2)	82 (42.5)	31 (16.1)	12 (6.2)	2 (1.0)	4.03	0.92
27	I like the way we learn writing in this class (e.g., hints before correction).	59 (30.6)	89 (46.1)	27 (14.0)	17 (8.8)	1 (0.5)	3.97	0.92
24*	I get bored in writing class because it doesn't hold my attention.	8 (4.2)	17 (8.8)	26 (13.5)	89 (46.1)	53 (27.5)	3.84	1.05

***Note:** Item 24 is a negative statement, reverse-coded—higher mean indicates lower agreement with "feeling bored."

Item 26 scored highest (M=4.30, SD=0.76), with 89.2% of students agreeing that DA activities were interesting. This largely stems from DA's collaborative elements: tasks like "group error-finding in sample essays" or "peer revision with teacher hints" turned passive writing practice into interactive experiences, unlike traditional classes where students write alone in silence (Wei, 2010).

Item 25 (M=4.03, SD=0.92) and Item 27 (M=3.97, SD=0.92) reflected moderate engagement: 76.7% of students found DA enjoyable, and 76.7% liked the learning methods. However, 16.1% (Item25) and 14.0% (Item27) were unsure, suggesting a subset of students still found elements of DA unengaging—likely low-proficiency learners who struggled with task difficulty (e.g., unable to participate in error-finding) and thus felt less interested.

Item 24 (reverse-coded, M=3.84, SD=1.05) had the lowest mean: 73.6% of students disagreed with "feeling bored," but 13.5% were unsure and 13% agreed. This uncertainty aligns with practical constraints: repeated revision activities (e.g., analyzing similar essay errors weekly) sometimes felt monotonous, even for engaged students.

Thematic analysis of open-ended responses (n=193) identified three core themes that explain quantitative trends:

Collaborative tasks boost engagement: 62% of students highlighted group activities as the most engaging part

of DA. A medium-proficiency student commented, "Working with classmates to find mistakes in a classmate's essay is fun—we compete a little to spot errors first, and the teacher praises the group that finds the most. It's way better than writing alone."

Realistic topics enhance interest: 48% of students mentioned that DA's focus on CET-4-related, relatable topics (e.g., "handling teacher-student relationships") made learning meaningful. A high-proficiency student noted, "Writing about topics we actually care about—like how to communicate with teachers—makes me want to put more effort in. In past classes, we wrote about boring topics like 'the importance of reading,' which I didn't care about."

Repetition and difficulty dampen interest: 29% of students (mostly low-proficiency) cited two pain points: repetitive revision tasks and task difficulty. A low-proficiency student stated, "After a few weeks, analyzing other people's essays feels the same—we look for the same mistakes. Also, when I can't find any errors, I get bored and zone out."

Kruskal-Wallis H tests revealed no significant differences in CAPTIVATE perceptions across high (M=4.11, SD=0.89), medium (M=4.01, SD=0.94), and low-proficiency (M=3.99, SD=0.93) groups (χ^2=2.151, df=2, p=0.341 > 0.05). This uniformity arises because DA's activity design balances engagement across levels: high-proficiency students enjoy the intellectual challenge of collaborative analysis, while low-proficiency students benefit from structured, low-pressure interactions (e.g., guided error-finding with hints)—both groups find value in the non-routine format.

The CAPTIVATE dimension confirms that DA makes writing instruction more engaging than traditional methods, with most students appreciating collaborative tasks and relatable topics. The lower scores in Item24 (reverse-coded) and Item27 highlight room for improvement: reducing repetitive activities (e.g., varying revision focuses weekly) and adapting tasks for low-proficiency learners (e.g., providing error checklists) would further boost engagement. Overall, CAPTIVATE demonstrates that DA's interactive, process-oriented design addresses a key pain point of college English writing—boredom—and lays the groundwork for sustained student participation.

Control: Perceptions of Classroom Order, Management, and Time Efficiency

Ranking seventh and being the only dimension with a mean below 4 (M=3.91, SD=1.02), CONTROL captures students' perceptions of how well teachers manage classroom order (e.g., reducing distractions), guide student behavior during DA activities, and utilize class time efficiently (Ferguson, 2010). This dimension ties closely to DA's unique challenge: unlike traditional writing classes—where students work independently and classroom management is relatively straightforward—DA's interactive, group-based activities (e.g., collaborative revision, peer discussion) require more flexible yet firm management to avoid chaos (Poehner, 2009). For Chinese college English classes with 40–50 students (Liu & Qi, 2024), this balance between interaction and order becomes even more critical.

Table 8 presents descriptive statistics for CONTROL items, including four reverse-coded negative items (Items 4–7) focused on student behavior issues. The results show notable variation, with high scores for respect and time efficiency, but low scores for overall behavior control.

Table 8. Descriptive Statistics for CONTROL Items

Item Number	Survey Item	SA (5) n (%)	A (4) n (%)	NS (3) n (%)	D (2) n (%)	SD (1) n (%)	Mean	SD
9	Students in this class treat the teacher with respect.	117 (60.6)	74 (38.3)	2 (1.0)	0 (0)	0 (0)	4.60	0.51
10	Our class stays busy and does not waste time.	97 (50.1)	80 (41.5)	11 (5.7)	2 (1.0)	3 (1.6)	4.38	0.77
8	My classmates behave the way my teacher wants them to during writing activities.	57 (29.6)	124 (64.3)	11 (5.7)	1 (0.5)	0 (0)	4.23	0.57

5*	I hate the way students behave in this class.	8 (4.2)	10 (5.2)	51 (26.4)	89 (46.1)	53 (27.5)	3.71	0.99
6*	Student behavior in this class makes the teacher angry.	7 (3.6)	11 (5.7)	61 (31.6)	71 (36.8)	43 (22.3)	3.68	1.00
7*	Student behavior in this class is a problem.	6 (3.1)	19 (9.8)	52 (26.9)	72 (37.3)	44 (22.8)	3.67	1.03
4	Student behavior in this class is under control.	31 (16.1)	45 (23.3)	44 (22.8)	57 (29.5)	16 (8.3)	3.09	1.22

***Note:** Items 5–7 are negative statements, reverse-coded—higher mean indicates lower agreement with "behavior issues."

Item 9 scored highest (M=4.60, SD=0.51), with 98.9% of students agreeing they respected the teacher. This reflects DA's positive teacher-student rapport (echoed in the CARE dimension): students' recognition of teachers' effort in mediation and feedback translated into mutual respect, reducing intentional disruptive behavior (Wallace & Chhuon, 2014).

Item 10 (M=4.38, SD=0.77) and Item 8 (M=4.23, SD=0.57) confirmed strong time management and general compliance: 91.6% of students felt class time was not wasted, and 93.9% reported classmates followed teacher guidance. DA's structured process—pre-writing → group revision → individual revision—kept activities on track, avoiding the "dead time" common in traditional classes (Wei, 2010).

Item 4 had the lowest mean (M=3.09, SD=1.22), with only 39.4% of students agreeing behavior was "under control"; 37.8% disagreed, and 22.8% were unsure. This stems from minor, frequent distractions (e.g., whispering during group work, phone use) in large classes—while not severe enough to disrupt learning, they made students perceive behavior as "not fully controlled."

Items 5–7 (reverse-coded, M=3.67–3.71) showed most students (69.9–74.1%) denied "serious behavior issues," but 26.4–31.6% were unsure—confirming that problems were mild but noticeable.

Thematic analysis of open-ended responses (n=193) identified three core themes that explain the quantitative split between "strong time/respect management" and "weak behavior control":

Structured DA process ensures time efficiency: 75% of students praised how tightly class time was organized. A high-proficiency student commented, "We never waste time—each part (writing, revision, summary) has a clear time limit. The teacher keeps us on track, so class feels productive, not messy."

Mutual respect reduces major disruptions: 63% of students linked respect to minimal serious behavior issues. A medium-proficiency student stated, "No one talks back to the teacher or skips activities—we know she's trying to help us improve. The worst that happens is someone whispers during group work, not big problems."

Minor distractions weaken "control" perception: 48% of students mentioned small disruptions that affected their perception. A low-proficiency student noted, "Some classmates play on their phones when we're revising essays. The teacher stops them when she sees it, but she can't watch everyone. It makes the class feel less controlled, even if it doesn't ruin my learning."

Kruskal-Wallis H tests showed no significant differences in CONTROL perceptions across high (M=4.00, SD=0.98), medium (M=3.86, SD=1.05), and low-proficiency (M=3.88, SD=1.03) groups (χ^2=3.526, df=2, p=0.172 > 0.05; Chapter 4-July 15.docx, Table 4.12). This uniformity arises because distractions and time management affect all students equally: high-proficiency learners were as annoyed by phone use as low-proficiency peers, and all valued the structured use of class time—regardless of their skill level.

The CONTROL dimension reveals a "mixed but positive" picture of DA classroom management. Students strongly confirm teachers' ability to utilize time efficiently and foster mutual respect, which offsets the impact of minor distractions. The low score for Item 4 ("behavior under control") highlights a practical area for improvement: targeted strategies (e.g., assigning "group leaders" to monitor small-group behavior, using quick

reminders for phone use) could address minor disruptions without sacrificing DA's interactive nature. Overall, CONTROL confirms that DA can maintain effective classroom order—even in large classes—by balancing structure with the collaborative needs of the approach.

DISCUSSION

This section interprets the empirical results of non-English major students' perceptions of Dynamic Assessment (DA) in college English writing instruction, connects the findings to the research questions, existing literature, and theoretical framework, and identifies the study's limitations and directions for future research.

Interpretation of Key Findings: Responding to Research Questions

The study set two core research questions to explore students' perceptions of DA and proficiency-based differences. The results provide clear answers to both questions, with distinct patterns reflecting DA's strengths and practical challenges in Chinese college English classrooms.

Overall Perceptions of DA Across Seven Dimensions

Research Question 1 focused on students' general perceptions of DA's seven dimensions (CARE, CONTROL, CLARIFY, CHALLENGE, CAPTIVATE, CONFER, CONSOLIDATE). The findings showed that students held predominantly positive attitudes (all dimensions $M > 3.9$), with three notable patterns:

Highest recognition for CONSOLIDATE (M=4.50) and CARE (M=4.48):This aligns with DA's core advantage of addressing the limitations of traditional static assessment. For CONSOLIDATE, students valued in-class summaries and peer revision—activities that helped them internalize writing knowledge (e.g., common errors, structural logic) and apply it to future tasks. This confirms Shrestha & Coffin's (2012) finding that DA's "feedback-integration" mechanism effectively bridges the gap between "receiving feedback" and "using feedback" in traditional writing instruction. For CARE, the high score reflects DA's ability to foster emotional rapport: teachers' individualized mediation (e.g., one-on-one guidance during revision) and attentiveness to writing anxiety (e.g., using Chinese to explain difficult concepts for low-proficiency students) reduced the emotional barriers non-English majors often face in writing (Zhang, 2023).

Lowest score for CONTROL (M=3.91)： The relatively low score, driven by Item 4 ("Student behavior is under control," M=3.09), highlights a practical challenge of DA in large classes. Unlike traditional classes where students work independently, DA's collaborative activities (e.g., group error-finding) occasionally led to minor distractions (whispering, phone use). While these disruptions did not undermine overall learning (Item 10, "Time efficiency," M=4.38), they made students perceive classroom order as "not fully controlled." This echoes Liu & Qi's (2024) observation that China's college English classes (typically 40–50 students) create inherent pressure for classroom management, even with DA's structured procedures.

Moderate scores for CAPTIVATE (M=4.03) and CONFER (M=4.12):These dimensions reflect DA's potential for improvement. For CAPTIVATE, 76.7% of students found DA enjoyable, but 14% felt unsure—largely due to repetitive revision tasks (e.g., analyzing similar essay errors weekly) and low-proficiency students' difficulty participating in error-finding. For CONFER, while 85.5% of students felt comfortable sharing ideas, 12% reported insufficient time to explain their thoughts, pointing to time constraints in 45-minute classes.

Proficiency-Based Differences in Perceptions

Research Question 2 examined whether students of high, medium, and low proficiency differed in their perceptions of DA. The key finding was that significant differences existed only in the CONFER dimension ($p=0.016$), with post-hoc tests showing high-proficiency students (M=4.26) scored significantly higher than medium-proficiency students (M=3.98), while low-proficiency students (M=4.12) did not differ significantly from either group. This pattern can be explained by DA's alignment with Vygotsky's Zone of Proximal Development (ZPD):

High-proficiency students:Their ZPD focused on "advanced expression needs" (e.g., proposing

counterarguments in essays). Teachers often adopted their suggestions (e.g., "Can we add a paragraph on solutions?"), reinforcing their sense of "being heard."

Medium-proficiency students:Their ZPD centered on "expression confidence." They were less likely to speak up in group activities, and when they did, their ideas (e.g., "Simplify this sentence") were sometimes overlooked in favor of high-proficiency students' more complex suggestions—leading to lower perceptions of "expression opportunities."

Low-proficiency students:Their ZPD prioritized "basic needs" (e.g., using Chinese to brainstorm). Teachers frequently accommodated these needs (e.g., allowing Chinese outlines), making them feel their voices were valued—hence no significant difference from high-proficiency peers.

This finding supplements existing DA research, which has rarely explored proficiency-based differences in perceptions (Kusumaningrum & Karma, 2018; Barnard, 2017). It challenges the implicit assumption that "all students perceive DA uniformly" and highlights the need for targeted mediation for medium-proficiency students.

Dialogue with Existing Literature

The study's results both validate and extend previous research on DA in L2 writing, contributing to the empirical base of DA's application in group classroom settings.

Validation of Existing DA Research

The findings align with three core conclusions of prior studies:

DA's effectiveness in integrating instruction and assessment: Consistent with Lantolf & Poehner (2004), the high scores for CONSOLIDATE and CLARIFY confirm that DA's contingent mediation (e.g., adaptive explanations, feedback integration) helps students internalize writing skills—unlike traditional assessment, which separates "testing" from "learning."

DA's potential to reduce writing anxiety: The high CARE score supports Ning's (2021) observation that DA's growth-oriented focus (e.g., emphasizing progress over grades) reduces non-English majors' writing anxiety. Students' open-ended comments (e.g., "I'm not scared to admit mistakes now") further corroborate this.

Group DA's feasibility in large classes: Despite CONTROL's low score, 91.6% of students reported efficient time use, and 89.2% found group activities (e.g., collaborative revision) engaging. This validates Poehner's (2009) claim that group DA can be adapted to large classes, as long as procedures are structured.

Extension of Existing DA Research

The study also adds two new insights to the literature:

Dimension-specific differences in DA perceptions: Most prior studies (e.g., Afshari et al., 2020) measured overall DA satisfaction, but this study's seven-dimensional analysis reveals nuanced patterns (e.g., CONSOLIDATE > CONTROL). This granularity helps identify DA's "strengths to leverage" (knowledge integration, emotional support) and "areas to optimize" (classroom management, activity variety).

CONFER as a proficiency-sensitive dimension：No previous study has identified CONFER as the only dimension with significant proficiency differences. This finding suggests that DA's "participatory" nature—often hailed as a strength—may inadvertently disadvantage medium-proficiency students, who need more scaffolding for idea expression (e.g., "expression templates" or "small-group pre-discussions") than currently provided.

Theoretical Implications: Aligning with Sociocultural Theory

The results strongly support Vygotsky's Sociocultural Theory (SCT), the study's theoretical framework, while

refining its application in DA-based writing instruction.

SCT's Core Tenets: Social Interaction and Internalization SCT posits that learning occurs through social interaction and the internalization of mediated support (Vygotsky, 1978). The high scores for CONSOLIDATE and CARE directly reflect this:

Internalization via CONSOLIDATE: In-class summaries and peer revision provided "social mediation" that students internalized—for example, high-proficiency students reported "actively avoiding subject-verb errors" after class summaries, while low-proficiency students used peer revision to "spot preposition mistakes in their own work."

Emotional mediation via CARE: SCT emphasizes that learning thrives in supportive social contexts. Teachers' individualized attention (e.g., taking off microphones during one-on-one guidance) created emotional safety, allowing students to engage with challenging tasks (e.g., revising complex paragraphs) within their ZPD.

Refining SCT: ZPD Variability Across Proficiency LevelsThe study extends SCT by highlighting ZPD's proficiency-specific nature：High-proficiency students' ZPD requires "advanced cognitive mediation" (e.g., discussing argument logic).

Medium-proficiency students' ZPD needs "expressive mediation" (e.g., building confidence to speak up).

Low-proficiency students' ZPD demands "basic linguistic mediation" (e.g., translating ideas from Chinese to English).

This refinement suggests that DA's mediation cannot be "one-size-fits-all"; instead, teachers must adapt support to each group's ZPD—an insight that enhances SCT's practical relevance for DA implementation.

LIMITATIONS OF THE STUDY

Despite its contributions, the study has three notable limitations that constrain the generalizability and depth of the findings:

Sample Limitations: The sample comprised only second-year accounting majors at Hebei Finance University (n=196). This narrow scope (single institution, single major) means the results may not apply to students in other regions (e.g., eastern coastal universities with higher English proficiency) or majors (e.g., humanities majors with more writing practice).

Methodological Limitations: Qualitative data relied solely on open-ended questionnaire comments, lacking triangulation with other sources (e.g., teacher interviews, classroom observations). For example, the low CONTROL score could not be fully explained—we do not know if it stemmed from teacher management strategies, student habits, or class size alone. Additionally, the 15-week intervention period may not capture long-term changes in perceptions (e.g., whether CAPTIVATE declines further with extended repetition).

Measurement Limitations: The modified Tripod Survey, while validated in Western contexts (Ferguson, 2010), may have cultural adaptation gaps. For example, the "CONTROL" dimension's focus on "student behavior" may be less relevant in Chinese classrooms, where respect for teachers (Item 9, M=4.60) already minimizes major disruptions—leading to potential misalignment between the survey and local contexts.

DIRECTIONS FOR FUTURE RESEARCH

To address the limitations and extend the study's findings, future research could focus on three areas:

Expanding Sample and Context: Future studies should include students from multiple institutions (e.g., rural vs. urban universities), majors (e.g., science vs. humanities), and proficiency levels to test the generalizability of the seven-dimensional perception pattern.

Longitudinal and Multimethod Design: A longer intervention (e.g., one academic year) could track how perceptions change over time (e.g., whether CONSOLIDATE remains high or declines). Incorporating teacher interviews and classroom observations would also help explain contextual factors (e.g., how teachers adjust mediation for medium-proficiency students in CONFER).

Developing Targeted DA Strategies: Based on the CONFER dimension's proficiency difference, future research could design and test "expression scaffolding" for medium-proficiency students (e.g., pre-writing idea-sharing worksheets, small-group leadership roles). Evaluating whether these strategies reduce perception gaps would provide practical guidance for DA implementation.

Culturally Adapting Measurement Tools: Refining the Tripod Survey to better align with Chinese classroom norms (e.g., redefining "CONTROL" to focus on "task engagement" rather than "behavior control") would improve the validity of perception measurements in local contexts.

In summary, the study's findings confirm DA's value in college English writing instruction while highlighting the need for targeted optimization. By addressing the limitations and pursuing the proposed future directions, researchers can further enhance DA's adaptability and effectiveness in Chinese higher education settings.

CONCLUSION

This study explores non-English major students' perceptions of Dynamic Assessment (DA) in college English writing and proficiency-based differences, aiming to inform teaching optimization in Chinese higher education. The key findings, implications, limitations, and future directions are briefly summarized as follows.

Core Findings

Students hold overall positive attitudes toward DA (all dimensions $M > 3.9$), with CONSOLIDATE ($M=4.50$) and CARE ($M=4.48$) most recognized for knowledge integration and emotional support. CONTROL ($M=3.91$) scored lowest due to minor distractions in large classes. Only the CONFER dimension showed significant proficiency differences ($p=0.016$): high-proficiency students ($M=4.26$) perceived more expression opportunities than medium-proficiency peers ($M=3.98$), while low-proficiency students ($M=4.12$) had no significant gaps.

Theoretical and Practical Implications

Theoretically, findings validate Vygotsky's Sociocultural Theory (SCT) and refine ZPD by highlighting proficiency-specific mediation needs. Practically, teachers should: (1) Reinforce in-class summaries and individualized support to leverage DA's strengths; (2) Adopt expression scaffolding (e.g., small-group pre-discussions) for medium-proficiency students; (3) Use group leadership or quick reminders to optimize classroom control.

Limitations and Future Directions

Limitations include narrow sample scope, single-source qualitative data, and potential cultural gaps in the measurement tool. Future research should expand samples, conduct longitudinal studies, design targeted DA strategies for medium-proficiency students, and adapt tools to Chinese classroom norms.

In short, DA is a valuable alternative to traditional assessment, effectively addressing writing instruction challenges. With targeted optimization for proficiency-specific needs, it can better enhance college English writing teaching quality.

REFERENCES

1. Afshari, M., Izadpanah, A., & Sadeghi, K. (2020). The impact of dynamic assessment on EFL learners' writing performance: A systematic review. Journal of Language Teaching and Research, 11(3), 456-465. https://doi.org/10.17507/jltr.1103.12
2. Barnard, R. (2017). Dynamic assessment in second language writing: A scoping review. Assessment in

Education: Principles, Policy & Practice, 24(5), 589-606. https://doi.org/10.1080/0969594X.2017.1327643

3. Ferguson, R. F. (2010). The Tripod Survey: Measuring teaching practices that matter for student achievement. Educational Evaluation and Policy Analysis, 32(3), 343-366. https://doi.org/10.3102/0162373710374709
4. Jacobs, H. L., Zinkgraf, S. A., Wormuth, D., Hartfiel, V. S., & Hughey, J. (1981). Testing ESL composition: A practical approach. TESOL Quarterly, 15(1), 31-43. https://doi.org/10.2307/3586451
5. Kusumaningrum, R., & Karma, D. (2018). EFL students' perceptions of dynamic assessment in writing classes. Journal of English Language Teaching and Learning, 3(2), 87-98.
6. Lantolf, J. P., & Poehner, M. E. (2004). Dynamic assessment in the language classroom: A teacher's guide. Modern Language Journal, 88(1), 1-13. https://doi.org/10.1111/j.1540-4781.2004.00324.x
7. Lantolf, J. P., & Poehner, M. E. (2009). Group dynamic assessment: Mediated development in the zone of proximal development. Language Teaching Research, 13(4), 445-463. https://doi.org/10.1177/1362168809342761
8. Liu, Y., & Qi, L. (2024). Classroom management challenges in large-scale college English classes in China. Journal of Higher Education in China, 7(1), 67-82.
9. Ning, Y. (2021). Alleviating EFL writing anxiety through dynamic assessment. System, 98, 102568. https://doi.org/10.1016/j.system.2021.102568
10. Poehner, M. E. (2009). Dynamic assessment: A Vygotskian approach to understanding and promoting second language development. Routledge.
11. Shrestha, R., & Coffin, C. (2012). Feedback integration in second language writing: The role of dynamic assessment. Journal of Second Language Writing, 21(2), 103-118. https://doi.org/10.1016/j.jslw.2012.03.002
12. Wallace, T. L., & Chhuon, V. (2014). Teacher-student rapport and classroom behavior: A cross-sectional study. Journal of Educational Psychology, 106(3), 789-801. https://doi.org/10.1037/a0036652
13. Wei, M. (2010). Time management in EFL writing classes: Comparing traditional and dynamic assessment approaches. Chinese Journal of Applied Linguistics, 33(4), 56-70.
14. Vygotsky, L. S. (1978). Interaction between learning and development. In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.), Mind and society (pp. 79-91). Harvard University Press.
15. Zhang, H. (2023). Individualized mediation in dynamic assessment: Supporting EFL learners' writing development. Journal of Asia TEFL, 20(2), 345-362. https://doi.org/10.18823/jate.2023.20.2.345