Effects of Hybrid Feedback Mode on College English Writing Teaching in China

Chuanlan Tang
School of Foreign Languages, Chongqing Jiaotong University, Chongqing 400000, China
tangchuanlan2022@163.com

Bing Li
School of Foreign Languages and Literature, Shan Dong University, Ji'nan Shandong 250100, China
libingcrystal@126.com

Heting Huang
School of Information Science, Beijing Language and Culture University, Beijing 100083, China
huangheting_blcu@163.com

Jiehua Zhong*
Faculty of Humanities and Social Sciences, Macao Polytechnic University, Macao 999078, China
p2212276@mpu.edu.mo

Abstract: The present study investigated the effects of hybrid feedback mode on Chinese college English writing teaching, including automatic feedback, peer+automatic feedback, self+automatic feedback and TSCA (teacher-student collaborative assessment)+automatic feedback. The interaction effect of students' English proficiency level, as well as their perception of the different feedback types were examined in detail as well. The participants were 270 sophomores from four classes of one university in Sichuan province in China. In order to investigate students' English proficiency level, an English proficiency test was employed to all students before the experiment, and it turned out that there was no significant difference ($F=0.640$, $p=0.59$) among the classes. In order to see the effect of different feedback methods, each class was employed different feedback types to evaluate their writing tasks over a period of 12 weeks experiment. All classes took the same pre- and post-tests, as well as 3 English writing tasks. In addition, at the end of the experiment, there were interviews and questionnaires conducted among the four classes. The results of this study suggested that all students' English writing grades had been improved significantly, besides, peer+automatic feedback and TSCA+automatic feedback had a more significant effect than self+automatic feedback and automatic feedback. However, the acceptance of TSCA+automatic feedback was relatively low, due to the inaccuracy and high demand of time. Thus, there should be more related research in the future. Meanwhile, there was an interaction effect between feedback types and

students' English proficiency level, which suggested that the proficiency level of students' should be taken into consideration in future research.
Keywords: Hybrid Feedback Mode; Automatic Feedback; Peer Feedback; Self-Feedback; TSCA; College English Writing

1. INTRODUCTION

1.1 Research Background

Writing is one of the key skills in language teaching, and some researchers believe that learners' success is related to their writing ability (Lerstorm, 1990). According to Pardiyono, writing is a process of sharing information, message (Pardiyono, 2007), or ideas in grammatically correct sentences. Writing emphasizes the appropriate use of vocabulary, syntax, and cohesive devices (Badger & White, 2000). Besides, writing has always been considered as an important skill in the teaching and learning of English as a Foreign Language (EFL) (Luchini, 2010).

Why is writing so important? First of all, it is not only because writing well—be it your own language, or a foreign one—clarifies your mind and orders your thoughts, but also because the ability to represent yourself well on paper will secure you a job in the future. Besides, according to a poll of 64 American companies, half of them evaluate writing when selecting a candidate for employment or promotion (Ariana, 2010). The level of writing is not only closely related to the number of words mastered and the application of the grammar but also related to the language application and expression ability of students (Wang, 2018). Improving English writing ability is way more important (Yu, 2020).

English writing is still the primary difficulty for most Chinese college students. According to the National CET-4 and CET-6 Examination Center: in June 1993, the average score of CET-4 writing was 6.8 points, with a scoring rate of 43.2%; In June 2001, the average score of CET-4 writing was only 7.7 points, and the scoring rate was only 51.3%, it can be seen that most students' English writing ability has not improved with the increase of English learning years. There are the following reasons for their poor English writing ability:

- Lack of enough English writing skills and knowledge: There are great differences in writing techniques between Chinese and English. However, many students do not understand the key differences between Chinese and English writing and often write in English according to the way they write in Chinese, which can affect the quality of their writing.
- Lack of qualified college English teachers: According to the latest

statistics, China has a population of 1.411178 billion, with more than 218 million college students. On the other hand, there are not enough college teachers according to the Ministry of Education of the People's Republic of China; In addition, faced with such a serious imbalance of teacher-student ratio, the quality of teaching must be affected to a certain extent. In the undergraduate colleges and universities that give priority to teaching, the proportion of professors and associate professors should reach 25%-35%. However, according to the data, only 19% of the 124 college teachers surveyed were professors and associate professors (Wang, 2014).

- Lack of effective feedback: With the increasing number of college students in China and the number of students in each class is exceeding the standard, it is difficult for teachers to give timely, comprehensive, and accurate feedback to each student. As a result, students tend to ignore the evaluation of writing and thus cannot effectively improve their English writing ability.

To sum up, at present, the English writing ability of Chinese college students is not only influenced by students' own ability and attitude but also closely related to teachers' teaching ability and feedback methods. Meanwhile, with the development of the times, especially the outbreak of COVID-19 in the past three years, there is a great urge for educators to seek new and more effective feedback methods with the trend of online education. As a result, this study, on the one hand, provides new ideas to Chinese college English writing teaching, on the other hand, it helps the practice of hybrid feedback model in English writing teaching.

2. LITERATURE REVIEW

Research on English writing feedback started very early in western countries. However, relevant studies in China started a bit late, especially the TSCA (teacher-student cooperation assessment) which was just put forward by Chinese education scholar Professor Wen in 2016. So the relevant theoretical and empirical studies are very limited.

After years of development, there have been many different versions of the automatic feedback system in China and abroad. For example, the earliest PEG system (Project Essay Grader) was developed by American professor Ellis Page in 1966; Intelligent Essay Assessor based on Pearson Knowledge Analysis Technology was published in the late 1990s, and the Electronic Essay Rater (E-Rater) extracurricular grading system is also very popular. The most popular automatic evaluation system in China is the

Juku system, Bingoenglish and so on. However, many scholars have different opinions on the function and effectiveness of automatic feedback system. For example, in the experiment of Wilson & Czik, compared with teacher feedback, students generally agree that automatic feedback is more time saving (Wilson & Czik, 2016). Through experiments, Li found that after using the automatic feedback system, students' compositions improved in many aspects (Li, 2017).

However, John & Woll believed that the automatic feedback system could not provide accurate and comprehensive feedback (John & Woll, 2020); Kim found through research that students were in urgent need of feedback at the sentence level, but the automatic system could not provide enough feedback (Kim, 2011). Peer feedback and self-feedback are complementary means of traditional teacher feedback. Peer feedback is a feedback method based on process teaching method, which is student-centered. There are also many forms of peer feedback, such as online peer feedback (forums, blogs, etc.) and written peer feedback (peer feedback checklist, etc.).

Self-feedback is a feedback type that students use their own essays as the source of feedback. There are many research on peer feedback and self-feedback. Berggren and Yang both believe that peer feedback can effectively improve students' writing performance (Berggren, 2013, 2015; Yang, 2017). On the other hand, Yang also found that only 38.7% of students liked to use peer feedback (Yang, 2017); Alhaisoni & Alzuoud believe that the effect of peer feedback depends on the prior training of students (Alhaisoni & Alzuoud, 2012). If students are trained properly, the effect of peer feedback can be enormous. Erdemir &Yesilanar found that Peer feedback had been criticized for being undetailed, inefficient, and subjective (Erdemir & Yesilanar, 2021). Sudarmaji &Lifanie found that among the three feedback methods: teacher feedback, peer feedback, and self-feedback, teacher feedback was the most effective one (Sudarmaji & Lifanie, 2020). Li mentioned that students preferred self-feedback than peer feedback (Li, 2017).

Teacher-student Cooperative Assessment (TSCA) was proposed by Professor Wen based on Production-oriented Approach (POA) (Wen, 2016). In the implementation process of TSCA, teachers only select some typical samples from students' writing and carry out targeted discussions in class, which includes three stages: before, during and after the class, and the actual operation is more complicated. Because the theory is quite new and difficult to operate, there are only few empirical studies on the theory.

Que believed that TSCA could significantly improve the accuracy, proficiency and enthusiasm of students' writing (Que, 2021). Sun believed that if students were familiar with the implementation process of TSCA (Sun, 2017), the cooperation between teachers and students could be better, as well as the effect. Moreover, the English proficiency of different students should also be considered when using TSCA.

In a word, most scholars only carried out independent empirical research on the automatic feedback system or only combine it with the traditional teacher feedback, and the research results are of great difference or even contradictory. In addition, the theory of teacher-student cooperative assessment (TSCA) was quite new, and there are only few related empirical studies, besides, most of them focus on its application alone instead of combining with automatic feedback.

Therefore, based on the actual situation of local colleges, this study combined peer feedback, self-feedback, and TSCA with automatic feedback respectively, to explore its application in practical teaching. At the same time, taking the differences of students' English proficiency level into full consideration, an empirical study is conducted on their possible interaction effects.

2.1 Research Significance

Alhaisoni & Alzuoud mentioned that in order for peer feedback to achieve great advantages, students need to receive correct training (Alhaisoni & Alzuoud, 2012). Therefore, this study trained students before the experiment, so as to make them more familiar with the process of peer+automatic feedback and to minimize the impact and errors caused by unfamiliar processes.

In the study of Yang, only 38.7% of students liked to use peer feedback (Yang, 2017). This study also investigated and analyzed students' acceptance of peer+automatic feedback. Sun mentioned that the implementation of TSCA in the future needs to consider the differences in students' English proficiency (Sun, 2017). This study conducted an English proficiency test for students, and analyzed the possible interaction effect between English proficiency level and different feedback modes.

To sum up, this study fully referred to the previous research, combined the automatic feedback mode with peer feedback, self-feedback, and TSCA. To make full use of the advantages of the different ways of feedback, try to reduce its own limitations, and strive to provide a new, more efficient hybrid feedback model for Chinese college English writing

teaching. Therefore, the research on effect of hybrid feedback mode on Chinese college English writing teaching is of great significance.

2.2 Research Purpose

This study explored the impact of hybrid feedback mode on students' English writing performance through empirical research, so as to seek a more efficient and adaptable feedback mode in new era. On this basis, this study also explored whether there was an interaction effect between students' English proficiency level and different feedback methods. After the experiment, interview and questionnaire survey were conducted to investigate the acceptance of students. The ultimate purpose of this study is to find an effective feedback mode for Chinese college English writing teaching, so as to improve the efficiency and level of college English writing teaching. In specific, the purpose of this research is to answer following questions:

(1) Are the writing scores of each group significantly improved?

(2) Are there any significant differences in the writing scores by different types of feedback?

(3) Is there any interaction effect between feedback types and students' language proficiency levels on their English writing ability?

(4) How do the students perceive their feedback method?

3. METHODOLOGY

Due to the complexity of the questions raised in the study, this study chose a combination of quantitative and qualitative research methods. The data collected during the experimental process, such as students' proficiency test scores and English writing scores, were quantitatively analyzed.

In order to understand the perception and acceptance of both teachers and students on the use of different feedback methods, this study also set up two sets of questionnaires, one for students participating in the experiment and the other for teachers with relevant teaching experience. At the same time, students were interviewed at the end of the experiments. The results of the interview was analyzed by the qualitative research method, in which Nvivo was used to conduct keyword analysis on the interview results and generate a word cloud map, which could more intuitively present students' perceptions.

3.1 Participants

There were 270 sophomores from four classes in a local university, located in Shichuan province in China, among them there were 119 female students and 151 male students. Their ages ranged from 19-20. The standard of the participants' selection was based on the last semester's English achievement task scores, and in order to ensure whether the participants are homogeneous or not, a proficiency level test was administered to the students before the experiment. The 270 students were divided into four groups: automatic feedback (Group 1), peer with automatic feedback (Group 2), self-with automatic feedback (Group 3), and teacher-student cooperative with automatic feedback (Group 4) (see Table 1).

Table 1: Participants' Profile

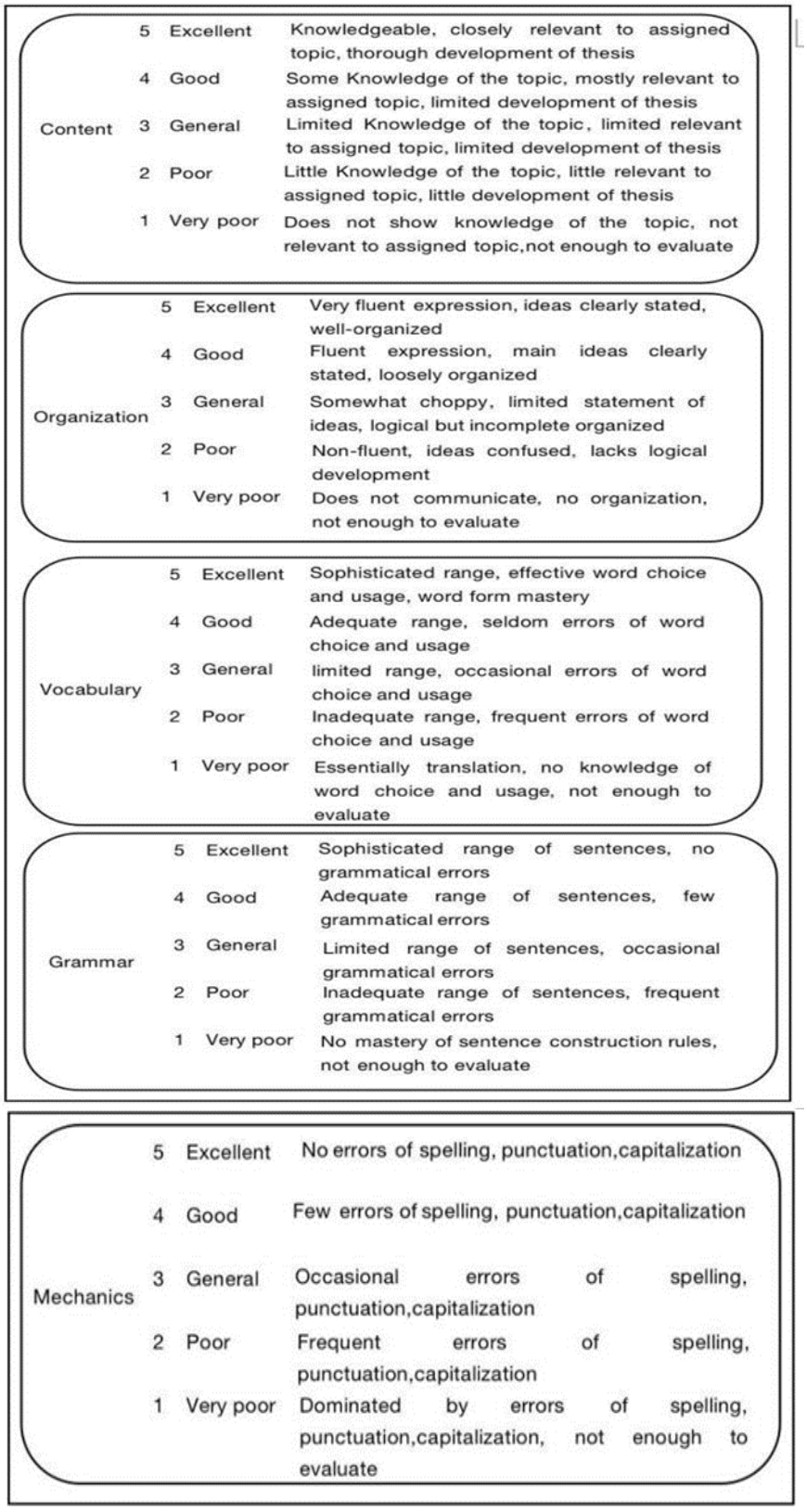
Group	Feedback Method	Grade	Age	Sex	Number	Total Number
1	Automatic feedback	Sophomore	19-20	Female	33	65
				Male	32	
2	Automatic+peer feedback			Female	27	67
				Male	40	
3	Automatic+self feedback			Female	31	70
				Male	39	
4	Automatic+TSCA			Female	28	68
				Male	40	
Total						270

3.2 Materials

- Proficiency test: Before the experiment, an English proficiency test was given to the students of the four classes to determine whether there were differences in English proficiency test scores in the four classes. The proficiency test consists of listening, reading, vocabulary, and grammar questions.

English writing tasks: In total, there are 5 English writing tasks (including pre- and post-test), and they all referred to CET-4 writing level.

Scoring criteria: Each piece of English writing was graded analytically. The total score of the essay was 25 points, which were divided into 5 points for each category: content, organization, vocabulary, grammar, and mechanics. The standard of scoring was based on Jacobs' scoring rubrics (see Figure 1) (Jacobs, 1981).



Content	5	Excellent	Knowledgeable, closely relevant to assigned topic, thorough development of thesis
	4	Good	Some Knowledge of the topic, mostly relevant to assigned topic, limited development of thesis
	3	General	Limited Knowledge of the topic, limited relevant to assigned topic, limited development of thesis
	2	Poor	Little Knowledge of the topic, little relevant to assigned topic, little development of thesis
	1	Very poor	Does not show knowledge of the topic, not relevant to assigned topic,not enough to evaluate
Organization	5	Excellent	Very fluent expression, ideas clearly stated, well-organized
	4	Good	Fluent expression, main ideas clearly stated, loosely organized
	3	General	Somewhat choppy, limited statement of ideas, logical but incomplete organized
	2	Poor	Non-fluent, ideas confused, lacks logical development
	1	Very poor	Does not communicate, no organization, not enough to evaluate
Vocabulary	5	Excellent	Sophisticated range, effective word choice and usage, word form mastery
	4	Good	Adequate range, seldom errors of word choice and usage
	3	General	limited range, occasional errors of word choice and usage
	2	Poor	Inadequate range, frequent errors of word choice and usage
	1	Very poor	Essentially translation, no knowledge of word choice and usage, not enough to evaluate
Grammar	5	Excellent	Sophisticated range of sentences, no grammatical errors
	4	Good	Adequate range of sentences, few grammatical errors
	3	General	Limited range of sentences, occasional grammatical errors
	2	Poor	Inadequate range of sentences, frequent grammatical errors
	1	Very poor	No mastery of sentence construction rules, not enough to evaluate
Mechanics	5	Excellent	No errors of spelling, punctuation,capitalization
	4	Good	Few errors of spelling, punctuation,capitalization
	3	General	Occasional errors of spelling, punctuation,capitalization
	2	Poor	Frequent errors of spelling, punctuation,capitalization
	1	Very poor	Dominated by errors of spelling, punctuation,capitalization, not enough to evaluate

Figure 1: Scoring Criteria of English Writing

- Juku system: Teachers assign a writing task on Juku in advance, including an essay topic, a writing requirement, word count requirement, a

passing grade, and completion time. The system automatically generates an assignment number, and the students can input this assignment number into the system to enter the writing page. For each submission, the system automatically scores the whole article and evaluate it sentence by sentence, including vocabulary, grammar, and discourse structure. Students can revise and submit the same essay numerous times within the prescribed time until they reach their own satisfaction. After the final submission, the teacher can see the total score of the student's essay, the times of submissions, the number of words, and the detailed comments on the essay.

- Peer feedback checklist: In order to measure the effectiveness of peer feedback, the group members were supposed to complete the peer feedback checklist (see Table 2). This checklist was based on the scoring criteria and referred to Wang, and some necessary modifications are made to ensure its efficiency (Wang, 2020).

Table 2(A): Peer Feedback Checklist

	Check items	**Modify Suggestion**
Vocabulary	Sophisticated range, effective word choice and usage, word form mastery ()	
	Adequate range, seldom errors of word choice and usage ()	
	limited range, occasional errors of word choice and usage ()	
	Inadequate range, frequent errors of word choice and usage ()	
	Essentially translation, no knowledge of word choice and usage, not enough to evaluate ()	
Grammar	Sophisticated range of sentences, no grammatical errors ()	
	Adequate range of sentences, few grammatical errors ( )	
	Limited range of sentences, occasional grammatical errors ()	
	Inadequate range of sentences, frequent grammatical errors ()	
	No mastery of sentence construction rules, not enough to evaluate ()	

Table 2(B): Peer Feedback Checklist

	Check items	Modify Suggestion
Organization	Very fluent expression, ideas clearly stated, well-organized ( )	
	Fluent expression, main ideas clearly stated, loosely organized ( )	
	Somewhat choppy, limited statement of ideas, logical but incomplete organized ( )	
	Non-fluent, ideas confused, lacks logical development ( )	
	Does not communicate, no organization, not enough to evaluate ( )	
Content	Knowledgeable, closely relevant to assigned topic, thorough development of thesis ( )	
	Some Knowledge of the topic, mostly relevant to assigned topic, limited development of thesis ( ) Limited	
	Knowledge of the topic, limited relevant to assigned topic, limited development of thesis ( )	
	Little Knowledge of the topic, little relevant to assigned topic, little development of thesis ( )	
	Does not show knowledge of the topic, not relevant to assigned topic,not enough to evaluate ()	
Mechanics	No errors of spelling, punctuation,capitalization ( ) Few errors of spelling, punctuation,capitalization ( )	
	Occasional errors of spelling, punctuation,capitalization ( )	
	Frequent errors of spelling, punctuation,capitalization ( )	
	Dominated by errors of spelling, punctuation, capitalization, not enough to evaluate ( )	

- Self-feedback checklist: In order to measure the effectiveness of self-feedback, group members were supposed to complete the self-feedback checklist (see Table 3). This checklist was based on the scoring criteria and some necessary modifications were made to ensure its efficiency.

Table 3: Self-feedback Checklist

	Check items
Vocabulary	Sophisticated range, effective word choice and usage, word form mastery ( ) Adequate range, seldom errors of word choice and usage ( ) limited range, occasional errors of word choice and usage ( ) Inadequate range, frequent errors of word choice and usage () Essentially translation, no knowledge of word choice and usage, not enough to evaluate ()
Grammar	Sophisticated range of sentences, no grammatical errors () Adequate range of sentences, few grammatical errors ( ) Limited range of sentences, occasional grammatical errors () Inadequate range of sentences, frequent grammatical errors ( ) No mastery of sentence construction rules, not enough to evaluate ( )
Organization	Very fluent expression, ideas clearly stated, well-organized ( ) Fluent expression, main ideas clearly stated, loosely organized ( ) Somewhat choppy, limited statement of ideas, logical but incomplete organized ( ) Non-fluent, ideas confused, lacks logical development () Does not communicate, no organization, not enough to evaluate ( )
Content	Knowledgeable, closely relevant to assigned topic,thorough development of thesis ( ) Some Knowledge of the topic, mostly relevant to assigned topic, limited development of thesis ( ) Limited Knowledge of the topic, limited relevant to assigned topic, limited development of thesis ( ) Little Knowledge of the topic, little relevant to assigned topic, little development of thesis ( ) Does not show knowledge of the topic, not relevant to assigned topic,not enough to evaluate ( )
Mechanics	No errors of spelling, punctuation,capitalization ( ) Few errors of spelling, punctuation,capitalization ( ) Occasional errors of spelling, punctuation,capitalization ( ) Frequent errors of spelling, punctuation,capitalization ( ) Dominated by errors of spelling, punctuation,capitalization, not enough to evaluate ( )

- Interview: At the end of the experiment, students in each group had a face-to-face interview lasting no more than 30 minutes. The interview was a semi-structured interview, which was designed to make clear subjective responses from people regarding a particular situation or phenomenon they had just experienced. Specific questions in the interview include:

1) Did you encounter any difficulties in using this feedback?
2) Can this feedback provide you with effective suggestions for improvement?
3) Do you think your English writing ability is improved through the use of this feedback?
4) Are you willing to continue using this feedback in future study?

- Questionnaires: In order to better understand the students' perception and acceptance of different feedback methods, a questionnaire survey was conducted to the students of the four groups at the end of the experiment.

The questionnaires were based on Ferris, and Xiang and some modifications were made (Ferris, 1995; Xiang, 2010). The questionnaire was written both in Chinese and English so that the students could understand it better. The questionnaire was designed by adopting a five-point Likert scale. And then different values were assigned: 5 for "Strongly agree", 4 for "Agree", 3 for "Not sure", 2 for "Disagree", and 1 for "Strongly disagree" for the choices of the answers. In total, there were fifteen questions in the questionnaire.

4.3 Procedures

The experiment began from March 7th to June 10th in 2022 and lasted for twelve weeks. The whole experiment process was closely supervised by the researcher to minimize interference factors. Before the experiment, the participating teachers were trained in English writing evaluation, the specific content of which was to learn the evaluation standards and scoring criteria of composition. The use of peer feedback checklist and self-feedback checklist was also explained to the relevant classes. The detailed procedure is presented in Table 4.

Table 4(a): Procedure of Experiment

Steps	**Activity**	**Teacher's Activity**	**Students' Activity**	**Time Limit**	**Goal**
1	Proficiency Test	Give 4 groups the English proficiency level test(including listening, reading and vocabulary) ; evaluate their test results.	Finish the proficiency level test on time.	60 min	To make sure all the groups have similar English level; to see if there are high level students and lower ones in each group.

Table 4(b): Procedure of Experiment

Steps	Activity	Teacher's Activity	Students' Activity	Time Limit	Goal
2	Pre-Test	Give 4 groups the English writing test; Evaluate their essays.	Accomplish the English writing test by themselves on time.	30 min	To make sure all the groups have similar English writing ability.
3	English Writing Task 1	Give students task 1	Accomplish the task on Juku and modify several times based on their own feedback.	1 month	To see if their feedback method improves their writing ability.
4	English Writing Task 2	Give students task 2	Accomplish the task on Juku and modify several times based on their own feedback.	1 month	To see if their feedback method improves their writing ability.
5	English Writing Task 3	Give students task 3	Accomplish the task on Juku and modify several times based on their own. feedback.	1 month	To see if their feedback method improves their writing ability.
6	Post-Test	Give 4 groups the English writing test and evaluate their essays.	Accomplish the English writing test by themselves on time.	30 min	To see if their feedback method improves their writing ability.
7	Interview & Questionnaires	Carry out interview among the students face to face and keep records; Ask students & teachers to finish questionnaires.	Participate in the interview & questionnaires.	60 min	To figure out their perception of different feedback methods.

4.4 Data Collection and Analysis

Scores of the proficiency test, pre-test, and post-test were recorded directly. Those three online writing tasks in the middle of the experiment were recorded through the automatic evaluation system (Juku),

supplemented by peer and self-feedback checklists. Interviews were organized in the form of interview records.The questionnaires at the end of the experiment were divided into valid questionnaires and invalid questionnaires, and invalid questionnaires were deleted from the result record.

Due to the diversity and complexity of experimental data and the large sample size, this study adopted the method of combining quantitative and qualitative analysis. Among them, the quantitative analysis used in this study included paired T-tests, independent sample T-tests, one-way ANOVAs, two-way ANOVAs, and post hoc analysis. Meanwhile, Nvivo was used for qualitative analysis. The proficiency test papers were automatically reviewed by the machine, and the results were analyzed and sorted through the descriptive analysis method, independent sample T-test, paired T-test, and one-way ANOVA.

In order to answer the first question (Are the writing scores of each group significantly improved?), this study collected the data of each group's writing scores on all writing tasks, including pre- and post-test and conducted paired T-test analysis among the groups.

In order to answer the second question (Are there any significant differences in the writing scores by different types of feedback?), this study analyzed and compared the writing scores of different groups with one-way ANOVA and post-hoc analysis. In order to answer the third question (Is there any interaction effect between feedback types and students' language proficiency levels on their English writing ability?), this study analyzed and compared the writing scores of pre-test and post-test of high- and lower-level students in each group. The results were analyzed using paired T-test and two-way ANOVA.

In order to answer the fourth question (How do the students perceive their writing feedback?), this study conducted interview and questionnaires among students and teachers respectively after the experiment and sorted out and analyzed the results with both quantitative and qualitative analysis methods.

5. RESULTS AND DISCUSSION

5.1 Pre-test Results

After the pre-test, the raters manually evaluated the students' compositions and sorted them out with descriptive analysis and a one-way ANOVA method. The specific results are shown in Table 5 and Table 6.

Table 5: The Descriptive Statistics of Pre-test

Group	N	Min	Max	Mean	SD	Range of Score (0-25)		
						<10	10-20	>20
Group 1	65	10	21	16.369	1.859	0	64	1
Group 2	67	11	20	16.149	2.134	0	67	0
Group 3	70	9	22	15.829	2.402	1	68	1
Group 4	68	10	19	15.559	2.622	0	68	0

As can be seen from the above table, the pre-test scores of students in the four groups are generally consistent, with the lowest score being about 10 and the highest score being about 20. Besides, in order to know if the four groups have similar English writing ability, one-way ANOVA was conducted among the pre-test scores (see Table 6).

Table 6: Results of Pre-test by One-way ANOVA

Item	Group								F	P
	Group 1 (n=65)		Group 2 (n=67)		Group 3 (n=70)		Group 4 (n=68)			
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		
Pre-Test	16.37	1.86	16.15	2.13	15.83	2.40	15.56	2.62	1.632	0.182

*p<0.05

As can be seen from the above table, through one-way ANOVA, the pre-test p value is 0.182 (p>0.05). Therefore, there was no significant difference in the English writing ability of students in the four groups before the experiment. This result, to a certain extent, ensures the consistency of the English writing level of the four groups before the experiment. It provides a foundation for the next English writing task experiment.

5.2 Post-test Results

At the end of the experiment, the English writing ability of the students in the four classes was post-tested and the results were sorted through descriptive analysis. The results are shown in Table 7.

Table 7: The Descriptive Statistics of Post-test

Group	N	Min	Max	Mean	SD	Range of Score (0-25)		
						<10	10-20	>20
Group 1	65	12	22	17.292	2.104	0	60	5
Group 2	67	15	24	19.463	2.488	0	42	25
Group 3	70	10	22	17.114	3.034	0	57	13
Group 4	68	12	25	18.515	3.414	0	46	20

As can be seen from the above table, after the experiment, the overall English writing level of the students in the four groups improved, among which the number of students with more than 20 points increased significantly, and even some with full marks. This result indicates that these different feedback methods have improved students' English writing ability significantly.

4.3 The Comparison of Pre- and Post-test

This section shows if there are significant differences in students' English writing scores in four classes before and after the experiment. Paired T-tests were adopted, and the results are shown in Table 8.

Table 8: The Comparison of Pre-test and Post-test

Item	Pairing				Difference (Pair 1- Pair 2)	T	P
	Pair 1		Pair 2				
	Mean	SD	Mean	SD			
Group 1	16.37	1.86	17.29	2.10	-0.92	-6.550	0.000*
Group 2	16.15	2.13	19.46	2.49	-3.31	-23.432	0.000*
Group 3	15.83	2.40	17.11	3.03	-1.29	-8.216	0.000*
Group 4	15.56	2.62	18.51	3.41	-2.96	-19.483	0.000*

* $p<0.05$

As can be seen from the above table, the paired T-test was used to study the difference in experimental data. There were four pairs of data in total, all of which presented significant differences ($p<0.05$). In a word, all the data of the four pairs in total show significant differences, which means that through the four different feedback methods, the English writing scores of these four groups of students have been significantly improved. Meanwhile, according to table, and comparing the data of the pre-and post-tests within these four groups, it can be concluded that Group 2 (-3.31) has the most significant difference, followed by Group 4 (-2.96), and Group 1 (-0.92) and Group 3 (-1.29) have no such significant difference. This also shows that automatic+peer feedback has the most significant effect, followed by automatic+TSCA feedback, and there is little difference between automatic feedback and automatic+self-feedback.

5.4 The Effects of Different Feedback Methods on Students' English Writing Scores

This section shows if there is a different degree of influence of different feedback methods on students' English writing performance. A one-way

ANOVA was adopted, and the analysis results are shown in Table 9.

Table 9: The Effects of Different Groups on Students' English Writing Scores

Item	Group									
	Group 1 (n=65)		Group 2 (n=67)		Group 3 (n=70)		Group 4 (n=68)		F	P
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		
Pre-Test	16.37	1.86	16.15	2.13	15.83	2.40	15.56	2.62	1.632	0.182
Post-Test	17.29	2.10	19.46	2.49	17.11	3.03	18.51	3.14	10.368	0.000*

*p<0.05

As can be seen from the above table, a one-way ANOVA was used to analyze the group's differences in pre- and post-test writing scores. The differences between the two items of writing scores can be seen in the above table as follows: Samples from different groups showed no significant difference in the pre-test writing score (p>0.05), which means that samples from different groups showed consistency in all the pre-tests writing score without any difference. However, there was a significant difference (p<0.05) among the samples in different groups in the post-test writing scores.

Different feedback methods used by the four groups have different effects on students' English writing ability, and the specific differences are analyzed by post-hoc differences, as shown in Table 10.

Table 10: The Post-hoc of Different Groups' Post-test

Group		Mean	Mean Difference	*p*	
Group 1	Group 2	17.292	19.463	-2.170	0.000*
Group 1	Group 3	17.292	17.114	0.178	0.714
Group 1	Group 4	17.292	18.515	-1.222	0.013*
Group 2	Group 3	19.463	17.114	2.348	0.000*
Group 2	Group 4	19.463	18.515	0.948	0.051
Group 3	Group 4	17.114	18.515	-1.400	0.004*

* p<0.05

According to the results of the post-hoc difference analysis, there are differences between Group 1 and Group 2 and Group 4. There are differences between Group 2 and Group 3. There are differences between Groups 3 and 4. And the comparative result of the group with a significant difference in average scores is: Group 2 > Group 1; Group 4 > Group 1;

Group 2 > Group 3; Group 4 > Group 3.

In conclusion, different groups of samples shows no significant difference in the pre-test writing performance, while different groups of samples shows a significant difference in post-test writing performance. In specific, on the base of usage of Juku, the influence of peer evaluation is greater than that of self-evaluation, and the influence of teacher-student cooperation evaluation is greater than that of self-evaluation as well. However, there is no significant difference between the influence of self-evaluation and automatic evaluation.

5.5 The Interaction Effect Between Students' English Proficiency and Feedback Types

5.5.1 Proficiency Test Results

After the proficiency test, the teacher reviewed the test papers and preliminarily sorted out the data with the descriptive analysis method. The results are shown in Table 11.

Table 11: The Descriptive Statistics of Proficiency Test

Group	**N**	**Min**	**Max**	**Mean**	**SD**	**Range of Score (0-60)**			
						<20	**20-30**	**31-41**	**≥42**
Group 1	65	17	49	39.031	6.177	1	3	39	22
Group 2	67	28	49	38.761	4.694	0	4	45	18
Group 3	70	28	48	39.014	4.985	0	7	40	23
Group 4	68	29	48	37.941	5.160	0	6	42	20

As can be seen from the above table, the lowest score of students in the proficiency test of the four groups is around 20, and the highest score is around 50.

In order to investigate is there any difference in students' English proficiency level of four groups, a one-way ANOVA was employed to analyze the results of proficiency test. The results shows in Table 12.

Table 12: The ANOVA Results of the English Proficiency Test

	Group (Mean±SD)					
	Group 1	Group 2	Group 3	Group 4	F	P
	(n=65)	(n=67)	(n=70)	(n=68)		
Proficiency Test	39.03±6.18	38.76±4.69	39.01±4.99	37.94±5.16	0.640	0.590

*p<0.05

From table above, it is revealed that there was no significant mean difference among the four groups' English proficiency level (F=0.640, p=0.590).

So all the groups were considered as identical groups. In order to investigate whether there is significant mean difference in high-level students and lower-level students in each group, independent sample T-tests were employed to each group. And the T-tests results shows in Table 13.

Table 13: The Independent Sample T-test Results of the English Proficiency Test of Each Group

Group	**Level**	**N**	**Proficiency level test**		**T**	***P***
			Mean	SD		
	High	22	44.91	2.20		
Group 1					9.473	0.000*
	Low	43	36.02	36.02		
	High	18	44.50	2.28		
Group 2					9.048	0.000*
	Low	49	36.65	3.40		
	High	23	44.39	1.85		
Group 3					11.946	0.000*
	Low	47	36.38	3.76		
	High	20	44.50	1.57		
Group 4					15.633	0.000*
	Low	48	35.21	3.32		

*p<0.05

From the table above, it can be seen that all high and low sub groups had a significant mean difference (Group 1: t=9.473, p=0.000; Group 2: t=9.048, p=0.000; Group 3: t=11.946, p=0.000; Group 4: t=15.633, p=0.000).

5.5.2 The Influence of Different Proficiency Levels on Students' English Writing Scores

This section shows if there are significant statistical differences in scores of students of different English levels in pre- and post-test. Paired T-tests were adopted, and the results were analyzed in Table 14.

Table 14: The Influence of Different Proficiency Levels on Students' English Writing Scores

Group	Level	N	Pre Test		Post Test		T	*P*
			Mean	SD	Mean	SD		
Group 1	High	22	18.18	0.50	19.36	1.40	-4.277	0.000*
	Low	43	15.44	1.59	16.23	1.54	-5.003	0.000*
Group 2	High	18	18.44	0.62	22.44	0.98	-15.646	0.000*
	Low	49	15.31	1.85	18.37	1.90	-19.696	0.000*
Group 3	High	23	18.39	0.66	20.61	1.03	-10.655	0.000*
	Low	47	14.57	1.89	15.40	2.05	-4.726	0.000*
Group 4	High	20	18.50	0.83	22.90	1.21	-22.295	0.000*
	Low	48	14.33	2.08	16.69	2.11	-20.090	0.000*

*p<0.05

As can be seen above, the paired T-tests are used to study the difference in experimental data. As can be seen from the above table, there are four pairs of data in total, all of which presented significant differences (p<0.05), indicating that the post-test scores of both high-level students and lower-level students in each group are significantly improved. In conclusion, both the high-level students and the lower-level students in the four groups have their English writing improved. Moreover, through the comparison of the high-level students and lower-level students in the four groups, it can be seen that the high-level students in each group made more progress than the lower-level students.

5.5.3 The Interaction Effect Between Students' English Proficiency Levels and Feedback Types.

This section shows if there is any interaction effect between students' own English level and different feedback types. A two-way ANOVA was adopted and, the results were analyzed in table 15.

Table 15: The Interaction Effect of Different Feedback Types and Proficiency Levels

Item	Sum of Squares	Df	The Mean Square	F	P	Eta2
Feedback types	284.047	3	94.682	31.661	0.000*	0.266
Proficiency levels	1237.751	1	1237.751	413.896	0.000*	0.612
Feedback types* proficiency levels	77.096	3	25.699	8.593	0.000*	0.090
Residual	783.507	262	2.990			

*p<0.05

As can be seen from the above table, different groups had significant mean differences on their post- writing test scores (F=31.661, p=0.000), indicating that the main effect existed, and different feedback types had a significant impact on the post-test writing scores. And the proficiency levels affected students' writing scores (F=413.896, p=0.000), indicating the existence of the main effect, and the proficiency level had a significant impact on the post-test writing scores. Besides, the result showed that there was an interaction effect between feedback types and proficiency levels (F=8.593, p=0.000). The high-level students can get more benefit from peer+automatic feedback and TSCA+automatic feedback than the other two feedback types. To be more specific, according to the Eta squared value, the influence caused by proficiency level accounted for 61.2%, which was larger than that of feedback types, indicating that the English proficiency level of students had a higher effect on their English writing.

5.5 Interview Results

This section presents the results of interviews among the students. At the end of the experiment, face-to-face interviews were conducted with the students of the four groups in order to understand the students' experiences and feelings about the feedback methods they used during the experiment. This study sorted out and selected 40 interview records for keyword analysis through Nvivo analysis software. The result is shown in Figures 2-4 below.

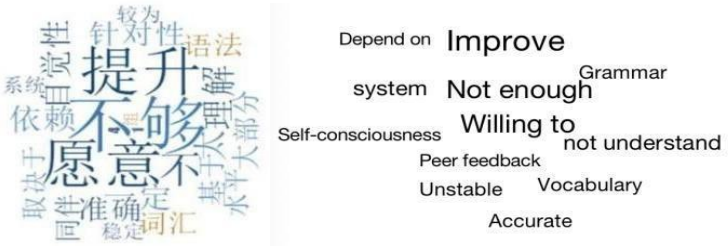


Figure 2: Key Words Cloud (Chinese & English versions)

单词	长度	计数	加权百分比 (%)
不够	2	49	12.96
愿意	2	40	10.58
提升	2	40	10.58
不	1	25	6.61
太	1	19	5.03
理解	2	19	5.03
一定	2	17	4.50
依赖	2	16	4.23
自觉性	3	16	4.23
准确	2	15	3.97
针对性	3	15	3.97
基于	2	13	3.44
大部分	3	13	3.44
词汇	2	13	3.44
语法	2	13	3.44
取决于	3	11	2.91
同伴	2	11	2.91
水平	2	11	2.91
稳定	2	6	1.59

The word	The length of the	count	Weighted percentage (%)12.96
Not enough	2	49	
Willing to	2	40	10.58
ascension	2	40	10.58
Not too	1	25	6.61
understand	1	19	5.03
A certain	2	19	5.03
Rely on	2	17	4.50
self-consciousness	3	16	4.23
accurate	2	16	4.23
targeted	3	15	3.97
Based on the	2	15	3.97
Most of the	3	13	3.44
vocabulary	2	13	3.44
grammar	2	13	3.44
Depending on the	3	13	3.44
companion	2	11	2.91
The level of	2	11	2.91
stable	2	11	2.911.59

Figure 3: Word Frequency of Students' Answers (Chinese & English versions)

访谈记录

访谈对象：第 2 组

访谈结构：半结构化访谈

访谈方式：小组访谈

面试时间：30 分钟

面试地点：706 教室

访谈内容：

问题 1：你在使用这个反馈时遇到过任何困难吗？
答：结果在很大程度上取决于你的同龄人的英语水平。有时，由于某些主观原因，合伙人不能提出客观、及时的修改意见。
问题 2：这个反馈能为您提供有效的改进建议吗？
答：是的。在大多数情况下，我们的合作伙伴可以提供及时和准确的修改意见。同时，结合自动评价系统的意见，我们可以对文章进行全面的改进。
问题 3：你认为使用这种反馈能提高你的英语写作能力吗？
答：是的。通过自动评价系统和同伴反馈的结合，对论文进行了多次修改，论文的最终分数与初审相比得到了提高。
问题 4：你愿意在未来的研究中继续使用这个反馈吗？
答：是的。这种反馈比纯机复习更准确，也可以促进学生之间的相互学习。通过指出你的伴侣的错误和缺点来提醒自己要进步。这样也可以提高我们的学习动机。

Interview Record

Interview object: Group 2

Interview structure: Semi-structured interview

Interview method: Group interview

Interview time: 30 mins

Interview place: Classroom 706

Interview content:

Question 1: Did you encounter any difficulties in using this feedback?
Answer: The results depend a lot on the English level of your peers. Sometimes, the partner cannot give objective and timely revision opinions for some subjective reasons.
Question 2: Can this feedback provide you with effective suggestions for improvement?
Answer: Yes. In most cases, our partners can give timely and accurate revision opinions. Meanwhile, we can improve the article comprehensively by combining with the comments given by the automatic evaluation system.
Question 3: Do you think your English writing ability is improved through the use of this feedback?
Answer: Yes. Through the combination of automatic evaluation system and peer feedback, the essay has been modified for many times, and the final score of the essay has been improved compared with the initial submission.
Question 4: Are you willing to continue using this feedback in future study?
Answer: Yes. This kind of feedback is more accurate than pure machine review, and can also promote mutual learning among students. Remind yourself to improve by pointing out your partners' mistakes and shortcomings. This way can also improve our learning motivation.

Figure 4: Interview Record (Chinese & English versions)

As can be seen from the figures above, students in Group 1 reflected

that the automatic feedback method was very dependent on the stability of the network and the system, and the evaluation of the content and structure of the article was not detailed enough. However, they agreed with the convenience and accuracy of this feedback method and expressed their willingness to continue to use it.

In the case of Group 2, students reflected that the automatic+peer feedback method was very dependent on the English level of their peers, which led to less accurate and comprehensive evaluation. However, the students also pointed out that this kind of feedback can promote the enthusiasm of students to learn from each other, which was more meaningful than using the automatic evaluation system alone, and expressed their willingness to continue to use it as well.

As for Group 3, students perceived that the automatic+self-feedback method was very dependent on their own English level and self-consciousness, which led to inaccurate and comprehensive evaluation. However, the students also pointed out that this feedback method was very convenient and could promote their self-examination ability to a certain extent and expressed their willingness to continue to use it.

Last, students in Group 4 reported that the automatic+TSCA feedback was much more difficult to conduct since the teacher only selected typical problems for class discussions, which left them still unsure of how to improve their own papers.

At the same time, this way of feedback needed to take up a lot of time and energy for both students and teachers. However, students also pointed out that this kind of feedback made them feel relieved because they could have the teacher's help. To recap the teachers' opinions and class discussions, TSCA+automatic feedback was more comprehensive than simply using an automatic evaluation system and the students expressed a willingness to continue to use it.

5.7 Questionnaires Results

At the end of the experiment, a questionnaire survey was conducted among the students of the four groups to understand their perception of the use of feedback. Group 1 received a total of 65 questionnaires, of which 60 were valid; Group 2 received a total of 67 questionnaires, of which 61 were valid; Group 3 received a total of 70 questionnaires, of which 64 were valid; Group 4 received a total of 68 questionnaires, of which 68 were valid. A one-way ANOVA was adopted, and the analysis results are shown in Table 16.

Table 16: The One-way ANOVA Results of Questionnaires

Item	Group								F	P
	Group 1 (n=60)		Group 2 (n=61)		Group 3 (n=64)		Group 4 (n=68)			
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		
Q1	3.73	0.94	3.98	0.90	3.91	0.95	3.99	0.94	0.988	0.399
Q2	3.70	1.00	4.03	0.84	3.92	0.91	3.99	0.89	1.594	0.191
Q3	3.83	1.04	4.15	0.85	3.98	0.92	4.10	0.93	1.375	0.251
Q4	3.83	1.04	4.30	0.72	3.95	0.95	4.13	0.88	3.070	0.028*
Q5	3.83	1.04	4.25	0.77	3.97	0.93	4.22	0.83	3.101	0.027*
Q6	3.93	0.71	4.28	0.88	3.83	0.97	4.18	0.83	3.778	0.011*
Q7	3.98	0.75	4.16	0.95	3.80	1.01	3.93	0.90	1.74	0.158
Q8	3.93	0.71	4.02	1.06	3.75	1.02	3.87	0.93	0.894	0.445
Q9	3.95	0.72	4.02	1.06	3.81	0.97	3.91	0.93	0.527	0.664
Q10	3.82	0.91	4.00	1.03	3.81	0.99	3.87	0.94	0.499	0.683
Q11	3.82	0.91	4.21	0.82	3.83	0.98	4.12	0.82	3.235	0.023*
Q12	3.93	0.73	4.23	0.88	3.80	1.04	4.15	0.85	3.142	0.026*
Q13	3.98	0.75	4.31	0.81	3.78	1.02	4.19	0.82	4.728	0.003*
Q14	3.92	0.72	4.16	0.92	3.77	1.02	4.21	0.84	3.598	0.014*
Q15	4.18	0.70	4.28	0.84	3.80	0.98	3.93	0.92	4.130	0.007*

*p<0.05

As can be seen from the above table, a one-way ANOVA was used to study the group differences of 15 items in Questions 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, and 15. As can be seen from the above table, different groups of samples do not show a significant difference (p>0.05) for Question 1, Question 2, Question 3, Question 7, Question 8, Question 9, and Question 10, which means that samples from different groups show consistency for Question 1, Question 2, Question 3, Question 7, Question 8, Question 9, and Question 10. However, the group samples showed significant differences for eight items in Questions 4, 5, 6, 11, 12, 13, 14, and 15 (p<0.05), which means different groups of samples have different responses to Questions 4, 5, 6, 11, 12, 13, 14, and 15. Specific analysis is as follows:

Different groups shows a significant level of 0.05 for Question 4 (F=3.070, p=0.028), and the specific comparison shows that the comparison result of the average score of four groups with a significant difference was 'Group 2>Group 1; Group 2>Group 3'; For Question 5,

the Groups presented significance at the level of 0.05 (F=3.101, p=0.027). As can be seen from the specific comparison differences, the comparison results of the average scores of the groups with significant differences are 'Group 2>Group 1; Group 4>Group 1'; The Groups shows a significant level of 0.05 for Question 6 (F=3.778, p=0.011). As can be seen from the specific comparison differences, the comparison results of the average scores of the groups with significant differences are 'Group 2>Group 1; Group 2>Group 3; Group 4>Group 3'; For Question 11, the Group presents significance at the level of 0.05 (F=3.235, p=0.023), and the specific comparison difference shows that the comparison results of the average score of the Group with significant difference are 'Group 2>Group 1; Group 2>Group 3'; For Question 12, the Group presents a significant level of 0.05 (F=3.142, p=0.026), and the specific comparison difference shows that the comparison results of the average score of the Group with a relatively significant difference are 'Group 2>Group 3; Group 4>Group 3'; The Group shows 0.05 level significance for Question 13 (F=4.728, p=0.003), and the specific comparison difference shows that the comparison results of the average score of the Group with an obvious difference are 'Group 2>Group 1; Group 2>Group 3; Group 4>Group 3'; For Question 14, the Group presents a significant level of 0.05 (F=3.598, p=0.014), and the specific comparison difference shows that the comparison results of the average score of the Group with a relatively significant difference are 'Group 2>Group 3; Group 4>Group 3'; The Group shows 0.05 level significance for Question 15 (F=4.130, p=0.007), and the specific comparison difference shows that the comparison results of the average score of the Group with an obvious difference are 'Group 1>Group 3; Group 2>Group 3; Group 2>Group 4'.

In conclusion, all four feedback methods can provide effective suggestions for mechanism, grammar, and vocabulary mistakes. Secondly, the feedback method of Group 2 and Group 4 can also give effective suggestions on the content adequacy, coherence, and structural logic of the essay and improve the writing level and enthusiasm of students to a greater extent. In addition, in terms of their willingness to use the feedback, students in Group 1 and Group 2 are more likely to use the feedback than those of Groups 3 and 4.

6. CONCLUSION AND SUGGESTIONS

For many Chinese students, writing has always been the most difficult part of learning English. In addition to the complexity of writing content, a big reason for their poor writing ability is that Chinese students lack

effective and comprehensive writing feedback.

This study conducted an empirical study on 270 students in four parallel classes over a period of three months. By using different writing feedback methods for four classes, this study explores the influence of different feedback methods on students' English writing ability and the perception and acceptance of students using different feedback methods. In view of the experimental data, this study adopts the method of combining quantitative and qualitative analysis. The purpose of this study is to provide theoretical and empirical support for relevant research fields and provide some new ideas and help for English teachers in future teaching as well.

6.1 Are the writing scores of each group significantly improved?

In order to answer this question, this study conducted pre- and post-test and three online English writing tasks for four classes. Through the analysis of writing scores, it could be seen that the English writing scores of students in each group had been significantly improved ($p<0.05$).

6.2 Are there any significant differences in the writing scores by different types of feedback?

In order to answer this question, this study employed a one-way ANOVA to analyze the data collected from post-test scores. The analysis results showed that there was a significant difference ($p<0.05$) among the samples in different groups in the post-test writing scores, which means that different feedback methods had different significant influences on students' English writing performance ($F=10.368$, $p=0.000$).

According to the results of the post-hoc difference analysis, the influence of peer+automatic feedback was greater than that of self+automatic feedback, and the influence of TSCA+automatic feedback was greater than that of self+automatic feedback as well. However, there was no significant mean differences between the group using self+automatic feedback and the group using automatic feedback.

6.3 Is there any interaction effect between feedback types and students' language proficiency level on their English writing ability?

In order to answer this question, this study submitted a two-way ANOVA to examine the interaction effect between feedback types and students' language proficiency levels. According to the analysis, there were statistically significant differences in the writing scores of high-level students and lower-level students in groups 1, 2, 3 and 4 ($p<0.05$),

indicating that the scores of each group's high-level students were better than those of lower-level students after the experiment. The interaction between feedback types and proficiency level showed a significant difference (F=8.593, p=0.000) as well, indicating that there was an interaction effect between the feedback types and language proficiency level. Moreover, the effect caused by proficiency level was larger than that of different feedback types.

6.4 How do the students perceive their writing feedback?

In order to answer this question, students of four classes were interviewed, and a questionnaire survey was conducted, Qualitative analysis software—Nvivo was used for the interview results, and one-way ANOVA and descriptive analysis were used for the questionnaire results.

According to the analysis, students stated that the automatic evaluation system had high requirements on the stability of the network and system, and the suggestions given by the system were not comprehensive and specific enough, but it was very convenient to use. Peer evaluation and self-evaluation depended on individual's English level, so that were somewhat subjective, but they could, to some degree, promote peer learning and improve self-learning ability. With the guidance from teachers, students who used teacher-students cooperative evaluation felt more secure. However, a class discussion could not deal with everyone's problems, which may lead to difficulty in modifying. At the same time, this feedback required much more time and energy from both students and teachers, so they were not willing to use it for a long time.

In addition, one of the limitations of this study is that the total sample size is limited. The research objects of this study are only four classes of students in one university, which may be not representative enough to reveal the situation. Another limitation is that the effects on different aspects of students' English writing have not been studied in depth due to the limited time. For example, content, organization, vocabulary, grammar, and mechanics of writing can be studied in detail, which may provide more specific support to English writing field.

Based on those limitations, there are some suggestions for future research. First, larger and more representative samples should be selected in future studies. Second, effects of different feedback types on different aspects of students' English writing should be studied in the future. Besides, since proficiency level can affect students' English writing scores in this study, there is a need to find out which feedback type is more helpful for high-level students and lower-level students separately in future study.

References

Alhaisoni, E., & Alzuoud, K. (2012). The effects of using peer and teacher feedback into Saudi writing context. *Indian Journal of Applied Linguistics*, *38*(2), 147-159.

Ariana, S. M. (2010). Some thoughts on writing skills. *Annals of the University of Oradea, Economic Science Series*, *19*(1), 127-133.

Badger, R., & White, G. (2000). Product, process and genre: Approaches to writing in EAP [Electronic version]. *ELT journal*, *54*(2), 153-160.

Berggren, J. (2013). Learning from giving feedback: Insights from EFL writing classrooms in a Swedish lower secondary school. *Stockholm University*.

Berggren, J. (2015). Learning from giving feedback: A study of secondary-level students. *ELT Journal 69*(1), 58-70.

Erdemir, N., & Yesilanar, S. (2021). Reflective practices in micro teaching from the perspective of preserver teachers: Teacher feedback, peer feedback and self-reflection. *Reflective Practice*, *22*(6), 766-781.

Ferris, D. (1995). Student reactions and teacher response in multiple draft composition classrooms. *Tesol quarterly*, *29*(1), 33-51.

Jacobs, H. L. (1981). Testing ESL composition: A practical approach. English composition program. *Newbury House Publishers*.

John, P., & Woll, N. (2020). Using Grammar Checkers in an ESL Context: An Investigation of Automatic Corrective Feedback. *Calico Journal*, *37*(2), 169-192.

Kim, T. E. (2011). Corrective feedback that an automatic writing evaluation system can and cannot provide. *English Teaching*, *66*(1), 111-140.

Lerstorm, A. (1990). Speaking across the curriculum; Moving toward shared responsibility? Paper Presented at the Annual Meeting of the Conference on College Composition and Communication, Chicago, IL.

Li, X. L. (2017). Self-feedback and peer feedback of students in oral English activities -- A case study of Liuzhou Vocational and Technical College freshmen. *Higher Education Conference*, *37*(4), 20-25.

Luchini, P. L. (2010). Evaluating the effectiveness of a complimentary approach to teaching writing skills. *International Journal of Language Studies (IJLS)*, *4*(3), 73-92.

Pardiyono. (2007). Teaching genre-based writing. *Yogyakarta: C.V Andioffset*.

Que, H. L. (2021). A practical study on the Teacher-student Cooperative Assessment (TSCA) created by POA in writing teaching for English majors. *University: Teaching and Education*(7), 43-44.

Sudarmaji, I., & Lifanie, M. B. (2020). A comparison study: Teacher, peer and self-feedback on students' essay writing. *Jisae: Journal of Indonesian Student Assessment and Evaluation*, *6*(2), 168-180.

Sun, S. G. (2017). The reflective practice of teacher-student cooperative evaluation in classroom. *Modern English*, *15*(3), 397-406.

Wang, D. D. (2014). Research on the current situation and development of college English teachers. *The Examination Weekly*, *35*(6), 30-32.

Wang, K. Y. (2018). The importance of high school English composition. *The Chinese Youth*, *6*(2), 16-21.

Wang, Y. X. (2020). A study on the influence of peer assessment on students' autonomous learning ability in high school English writing teaching. *Master dissertation, Qufu Normal University*.

Wen, Q. F. (2016). "Evaluation of Teacher-student Cooperation ": A new evaluation form created by" Output Oriented Method". *Foreign Languages*(5), 37-43.

Wilson, J., & Czik, A. (2016). Automated essay evaluation software in English language arts classrooms: Effects on teacher feedback, student motivation and writing quality. *Computers & education*, *100*(2), 94-109.

Xiang, Y. (2010). The role of teacher feedback and peer feedback in college English writing. *Master dissertation. Chongqing University*.

Yang, J. D. (2017). A comparative study of teacher feedback and peer feedback in college English writing. *Journal of Ningbo Institute of Education*, *19*(1), 66-70.

Yu, L. J. (2020). The importance and problems of English writing. *Writing Guidance for Middle School Students*, *10*(4), 2-5.