Vol.3 No.3 2022
https://doi.org/10.37420/j.cer.2022.028

Research on Teaching Assessment Mode Of English For College Based on Intelligent Language Diagnostic Test

Xiaoyan Zheng*
Kunming Railway Vocational and Technical College, Yunnan, China
*Corresponding Author, e-mail: 3240686169@qq. com

Abstract: *Online testing, traditional testing and mixed testing mode are adopted in the teaching procedure and final assessment in English for Colleges. For the English learning cognitive and characteristics of vocational college students, diagnostic tests combined with advanced intelligent technology appeal to apply into wide and real practice. This research takes Grade One students majoring in Urban Rail Vehicles, Railway Engineering, and EMU Maintenance Technology as samples. 8 to 10 students whose English scores were in mediate and low level in each class are high-lightly traced. Intelligent diagnostic tests in English listening comprehension, grammar and translation are conducted for the samples in Chaoxing Xuexitong. By interviews and diagnostic tests, the research group continuously traces on the samples' English application skills, quickly positions the samples' learning level in Listening & Speaking, Grammar and Translation, adjusts the teaching approaches, reconstructs the English assessment system, and sheds the light for high quality development of language teaching and learning.*

Key words: *diagnostic tests, intelligent, teaching assessment, English for college.*

Introduction

In recent years, the developing speed and application of artificial intelligence technology are far beyond people's expectation. Especially at the beginning of 2020, since the outbreak of the covide epidemic, online courses or on-off line mixed courses have become the new but normal mode. On the other hand, the traditional testing mode with the paper and pencil has been gradually transferred to the online as well. In this phase, the commercialization operation of language testing, diagnostic testing, testing of a foreign language for special purposes have been gradually conducted into practice to meet the diversified social needs. Online learning softwares such as Unipus, Xuexitong, etc are widely used in teaching. Scholars have also carried out practice in combination with teaching, performing the teaching design of diagnostic language test and creating intelligent classroom. For example, under the guidance of communicative language test theory, English test in oral speaking for vocational colleges designed in layers (Chen & Yang 2021; Zhu, Dan 2020); AI enpowered in college English Writing and Reading (Du 1999; Du&Gui 2000; He 2011); Language testing technology widely used in CET-4 and CET-6 to comprehensively assess English listening, speaking, reading, writing and translation (Chen 2009; Jin 2022). The increasing decision-making role, "Internet +" and the enhancement of students' demand for independent learning provides favorable conditions for intelligent diagnostic testing. Some scholars believe that although "Internet +" has been widely applied in education from top to bottom and from point to surface in an unprecedented way, intelligent language testing still has a series of problems : (1) the application of technology may show the representativeness of test constructs (Li 1998; Jin 2011&2020),

for example, in order to achieve automatic scoring, test task design may try to adapt to the requirements of the automatic scoring system, such as controlling the openness of examinees' answers; (2) The evaluation of technology application should focus on whether the test is beneficial to promote teaching and students' independent learning. For example, students deliberately seek and make use of the so-called "game rules" of the scoring system by showing the characteristics that the system pays attention to, such as the frequent use of long sentences or low-frequency words to obtain high scores; (3) The materials of native speakers are still used as input materials for testing tasks, and the performance of native speakers is still used as the reference standard for scoring (Jin& Sun 2020). (4) Students' learning is a gradual process, and a comprehensive understandingt requires continuity of diagnosis; otherwise, only fragmentary information can be obtained.

Theory For Language Assessment Test

Bacheman (1990) and Shohamy (2017) proposed that language testing should always focus on test constructs, assessing methods and the sociological dimensions, namely what to test, how to test and for what. Jin (2020) proposed that Shohamy, in his book Language Testing and Assessment, had introduced and solved the research directions in the field of language testing, including (1)assessing language domains, such as the test of meaning, the test of multilingual competence, and the test of English as a lingual language. (2) methods of assessment; (3) assessment in education, such as test effectiveness and assessing literacy; (4) assessment in society, such as the relationship between high-risk tests and language education policy, as well as the professional ethics and behavior standard in the field of language test. Diagnosis refers to the act of discovering or identigying the exact cause of an illness or a problem (Oxford Dictionary). Diagnostic test in college English teaching refers to the examination of the weak links in the learning process of students, so as to make up for the past shortcomings, adjust the teaching plan and improve the quality of classroom teaching. Diagnostic tests are phased tests, including unit tests, midterm tests, or final tests. The content of the test is limited to the teaching content of the period, and the test can be in various forms, including formal examination questions, filling in blanks, completing sentences, spelling words, etc (Huang 2021).

Design for Research*

Online diagnostic tests in an open ways, regular interviews and online questionnaires are conducted by the research group. Grade one students majoring Urban Rail Vehicles, Railway Engineering,and EMU Maintenance Technology as samples, 8 to10 students whose English is in intermmediate and lower level in each class are highlighted (Class A: 10 Ss, Class B: 7 Ss, Class C: 8 Ss, a total of 25). Intermediate students are defined as having a total score of 65-75 on the final assessment at the first semester (Examing Scheme seen Figure 1), while lower students have a score of less than 64. In the second semester, intelligent diagnostic tests are performed in English listening, grammar and translation. The diagnostic test is conducted in the second semester so that there could be be an objective comparasion with the first semester in terms of the the final assessment scores. The design of the test questions is based on the Level A Practical English Test. Diagnostic tests are proceeded at the end of each unit, while interviews are planned to be done at the end of Unit 2 and Unit 5, and online questionnaires on course assessemt are performed at the end of Unit 8(seen Figure 2). Through the continuously tracking of English listening and speaking, grammar and translatiofor intermediate and lower students. The test data and psychological cognitive changes of samples are analyzed, and the teaching strategies are adjusted in time to act on the remedial teaching directly.

Vol.3 No.3 2022
https://doi.org/10.37420/j.cer.2022.028

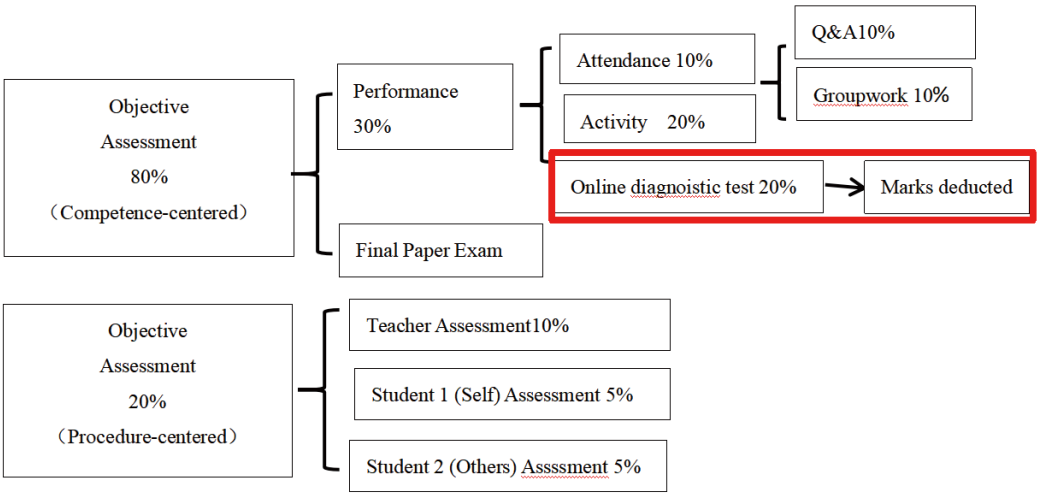

Figure 1 Assessment Mode of Vocational English Teaching With Online Diagnoistic test Added in Semeser II

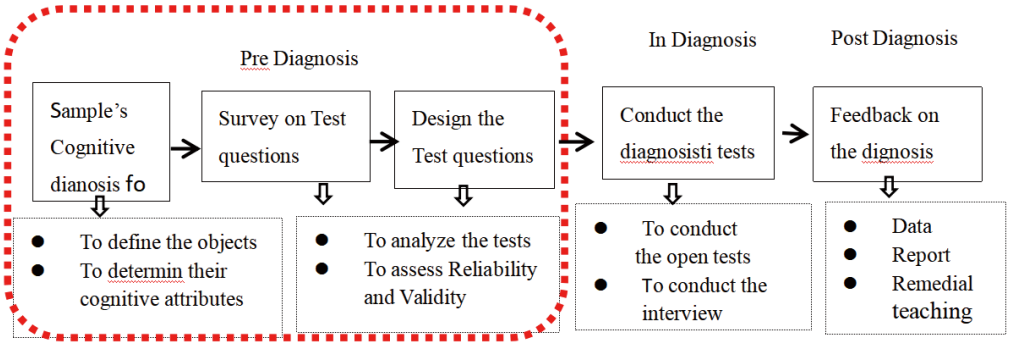

Figure 2 Implementation Path of Intelligent Language Diagnostic Test

Selection of test samples and carriers

Samples

Taking the Grade One students in author's school majoring in Urban Rail Vehicles, Railway Engineering,and EMU Maintenance Technology as samples, 8 to10 students whose English is in intermmediate and lower level in each class are highlighted, among those, 10 students in 21 Urban Rail Transit (Class A), 7 in 21 Railway Engineering(Class B) and 8 in 21 EMU Maintenance Technology(Class C). There are 25 students in total (2 of them are Japanese examinees in Gaokao). 1 female, accounting for 4%; 33 Male, accounting for 96%.

Online testing carriers

21st Century Practical College English series, edited by Yu Jianzhong, Wang Chaohui and Liu Shuang, is currently being used by the research group. The digital course is based on the teaching materials and completely presents the teaching materials. According to the rules of students' online learning, Xuexitong further

develops online learning links such as oral exercises, following reading and audio assessment to help students make full use of fragmented time to learn, develop their independent learning habits and improve their independent learning ability. By providing functional modules matching the corresponding scenes for assistance in each link of Pre-class, In-class, After-class, it offers good digital support for teachers' mixed teaching and students' independent learning. In addition to support the teaching process, Xuexitong can also assist in the formative assessment of students. Based on formative assessment, teachers can trace the teaching effect from two demensions: class and students. The intelligent system will automatically analyze the students' course learning in details from each module; teachers can view each student's learning data such as learning proceed, scoring, learning length in the student details. Teachers can timely offer intervention and guidance to students with difficulty, The above data can be viewed or exported on the mobile phone.

Survey and construction on test question bank

The questions of test takes the Level A Practical English Test as the reference, and are set as objective ones with choice and fill-in-the-blank, including listening, translation and grammar. According to the survey on the10 sets of Level A Practical English Test from 2016 to 2021, it is found that the focus of tests is fixed collocation(V+O; P+O) accounting for 16%, Non-finite verbs 14%, tense 11%, adverbial clauses 10%, nouns and gerunds 9%, comparative adjective and attributive clauses 6%, adverbs, inverted sentences, emphatic sentences and subjunctive mood 4%, infinitive and subject clauses 2%, conjunctions, synonym analysis and object clauses 1%(seen Figure 3). Translation in business is dominant (customer service); covered employment, technology, the environment (seen Figure 4). The diagnostic test in Xuexitong is designed with objective questions and marked online in real time. The online test contains grammar, translation and listening, with the weight of points accounting for 50%, 30% and 20% respectively. The test points are the same as the Level A Practical English Test.

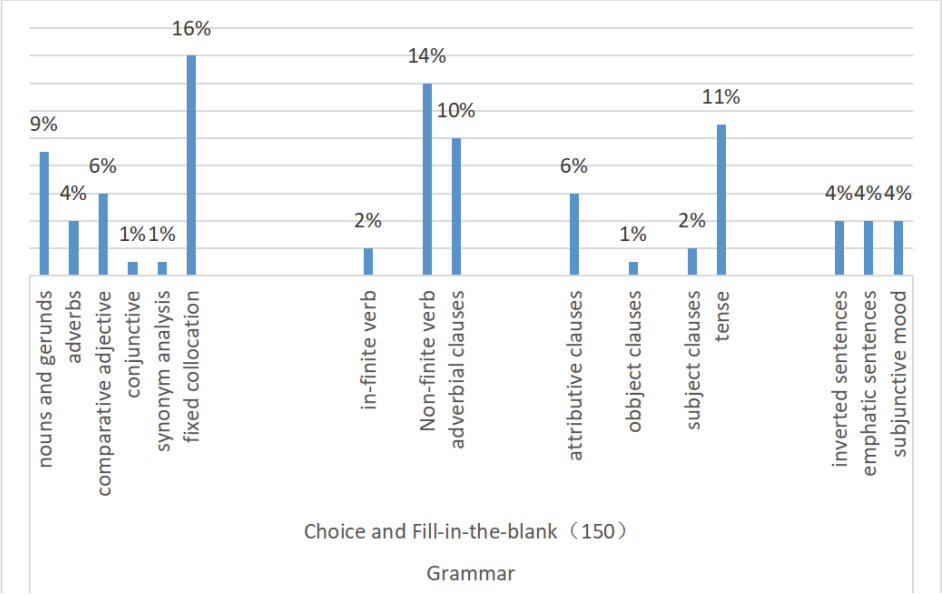

Figure 3 Test Points of Grammar in Level A Practical English Test from 2016 to 2021

Vol.3 No.3 2022
https://doi.org/10.37420/j.cer.2022.028

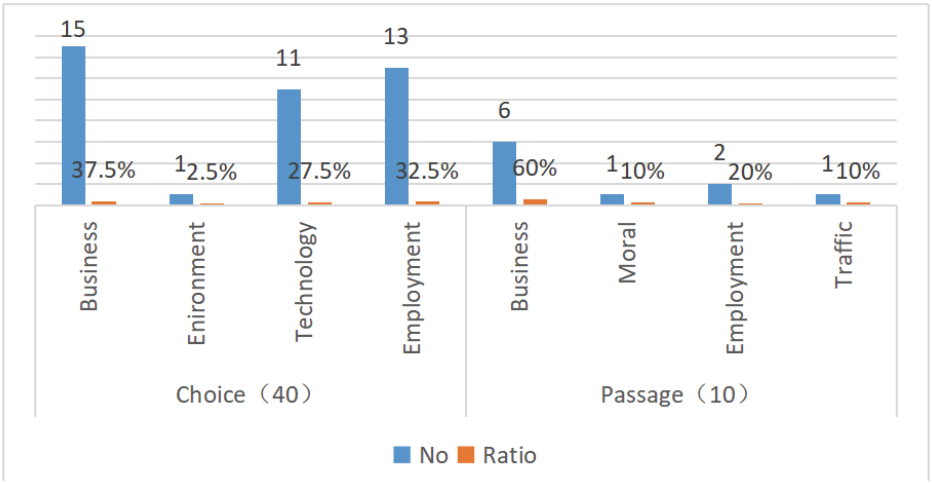

Figure 4 Test Focus of Translation in Level A Practical English Test from 2016 to 2021

Design the test questions

The design principle of the test questions is to make better use of the student's fragmented tim, seek the weak points of knowledge, promote the "deep learning" of students, and realize "quantitative change" in learning time to the "qualitative change" in learning quality and independent learning ability. With the help of data and analysis of the platform, teachers can reflect on teaching and empower the sciencetific research. While conducting the test questions, considering the effective promotion of students' English application ability and independent learning ability, it is required that 30% the test comes from coursebooks, 30% from Level A Practical English Test, and 40% the extracurricular. The research group takes A-level examination questions as the reference for the design of the test question bank, and the tested points in each module of Level A Practical English Test as the domaint (see Figure 5). The reason why reading is not included in the designed test is that students generally believe that listening, grammar and translation are the difficulties in English learning, and their reading ability is relatively good. Meanwhile, a long-time online answeingr will also have adverse effects on students' attention and eyesight, so the testing time is set in 60 minutes. However, student can repeatly do the test in the regulated time, and the highest score in each single test will be recorded in the intelligent system.

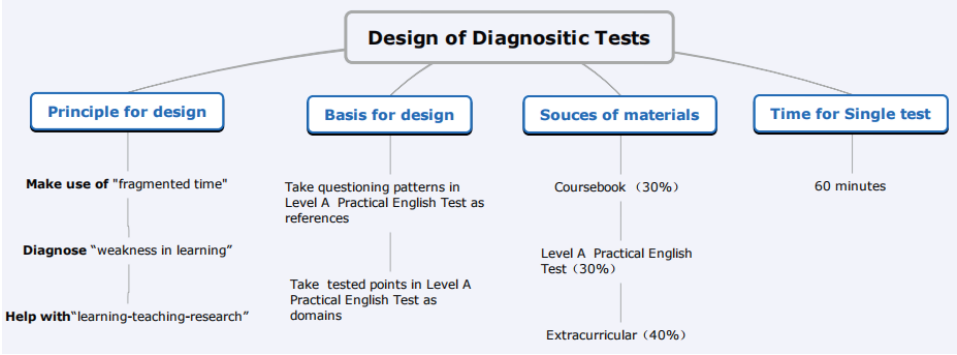

Figure 5 The Design of Diagnostic Tests

Take the Listening in Diagnostic Test 8 as an example:

• **Section A**

Click on audio clip

Directions: *This section is to test your ability to understand short dialogues. There are 5 recorded dialogues in it. After each dialogue, there is a recorded question. When you hear a question, you should decide on the correct answer from the 4 choices marked A, B, C and D given.*

Dialogue 1 (From extracurricular)
Female: How do I develop the practice the of gratitude?
Male: You can list three to ten things which you are grateful. You can do this before you go to bed every night.
Question; What is the man's suggestion about?
A. How to make one's daily life easy.
B. How to write a letter of gratitude.
C. How to focus on the meaning things.
D. How to develop the practice of gratitude.

• *Illustration:* It tests the students' understanding for the word *gratitude* and the sentence. The topic about gratitude appeals to THANKS in Unit 3(Book I). D is the correct choice.

• **Section B（From extracurricular）**

Click on audio clip

Directions: *This section is to test your ability to understand short conversations. After each conversation, there are some recorded questions. When you hear a question, you should decide on the correct answer from the 4 choices marked A, B, C and D given.*

Male: I can see from your resume that you like meeting people and enjoy making new friends.
Female: Yes. I am that sort of person.
Male: Well. That is exactly the kind of person we need.
Female: Believe me I can do the job well.
Male: But you want to be meeting the people and person only over the phone.
Female: I have a good telephone voice. And i am very confident on the phone.
Male: Ok. We need someone for 3 evenings a week at least 3 hours an evening.
Female: No problem. I can do that.
Male: That fine. Please come with me. And I will introduce you to the manager.
Female: Thank you very much.
（105words）

Vol.3 No.3 2022
https://doi.org/10.37420/j.cer.2022.028

Question1: What can we learn about the woman from the conversation?
A. She enjoys making new friends.
B. She likes doing volunteer work.
C. She gets along well with her colleagues
D. She possesses the qualities of a good leader.

• *Illustration:* It tests the students' listening and understanding for details. Here is the clue in *I can see from your resume that you like meeting people and enjoy making new friends.* A is the correct choice.

Question2: What kind of the job is the woman going to take?
A. Meeting people in person.
B. Talking with people over the phone.
C. Introducing products to customers.
D. Making schedules for the manager.

• Illustration: It tests the students' listening and understanding for details. *Here is the clue in But you want to be meeting the people and person only over the phone*. B is the correct choice.

Question3: How many hours will the woman work a week?
A. Three. B. six. C. Nine. D. Twelve.

• Illustration: The question is much difficult, requiring the understanding and quick calculation. Here is the clue in *3 evenings a week at least 3 hours an evening.* C is the correct choice.

• **Section C**（*From Coursebook Unit 1 Book II*）

Click on audio clip

Directions: In this section you will hear a recorded short passage. The passage is printed in the test paper, but with some words or phrases missing. You are required to put the missing words or phrases on the Answer Sheet in order of the numbered blanks according to what you hear.

In our daily conversations, we cannot help asking for clarification when we are not clear about what (1)has been said or have become puzzled about a particular point (2)mentioned earlier. A conversation is a form of two-way interaction. During interaction, misunderstanding can lead to (3) a breakdown in communication. To avoid misunderstanding or (4)even confusion, we tend to have an impulse to ask for clarification. This is especially true when a Chinese learner of English talks with a native speaker. This skill can help him or her avoid embarrassment caused by misunderstanding, and keep the conversation (5)going on. (97words)

• *Illustration:* It is about language misunderstanding and clarification, which provides correct guidance to students' English learning. It tests the students' understanding of the passage and spelling of Level A vocabularies.

• **Section D**（*From extracurricular Coursebook Unit 7 Book II*）

Click on audio clip

Directions: This section is to test your ability to comprehend short passages. You will hear a recorded passage. After that you will hear five questions. When you hear a question, you should complete the answer to it with a word or a short phrase (in no more than 3 words).

We've all heard the classic adage (格言), "Is the glass half empty or half full?" but we may not realize the true power of seeing the world through rose-colored glasses. It can have a far-reaching (深远的) influence on everything from our day-to-day lives and selections of mates (配偶), to our career choices and long-term success. Optimists think in color. They are open-minded to new ideas and situations and enjoy the thoughts and opinions of others. Optimists realize that change is a good and necessary thing. They usually maintain (保持) an altitude of adaptability (适应) and take criticism with a grain of salt (有保留地). They always, however, appreciate good feedback (反馈) and constantly look to bettering themselves. Learning optimism makes more enjoyable life. (115words) .

Q1: What is the classic adage relevant to the talk?
A: Is the glass half empty or half full.
Q2: What is the power of seeing the world through rose-colored glasses?
A: It can have a far-reaching influence on everything
Q3: What is optimist thinking style?
A: They are open-minded to new ideas and situations and enjoy the thoughts and opinions of others.
Q4: What do optimists think of change?
A: They think that change is a good and necessary thing.
Q5: How do they react to feedback and take criticism?
A: They appreciate good feedback and take criticism with a grain of salt.

• *Illustration:* The topic is Optimism(乐观主义) which conducts a positive attitude. It tests the students' understanding of the passage and spelling of Level A vocabularies.

Outcome and Anaysis

Tracking on Learning independence of test samples

Through the diagnostic test, the independent learning initiative of the samples is significantly enhanced. The specific performance is as follows: ① The study time for each sample becomes at least 1 hour a day from less than 30 minutes at the begining; students highlight the tests, for the tests have been done more than 10 times for each student (seen Figure 6 ; Data source: Xuexitong, June 31, 2022); ② Diagnostic tests guide the samples's behavior to conduct self-assessment (seen Figure 7); ③ The "deep learning" awareness of the samples has begun to awaken, and they tend to come up with effective learning strategies, such as actively taking notes in class, asking teachers and classmates for help besides seeking online for answers, checking the correctness of the language output.

Vol.3 No.3 2022
https://doi.org/10.37420/j.cer.2022.028

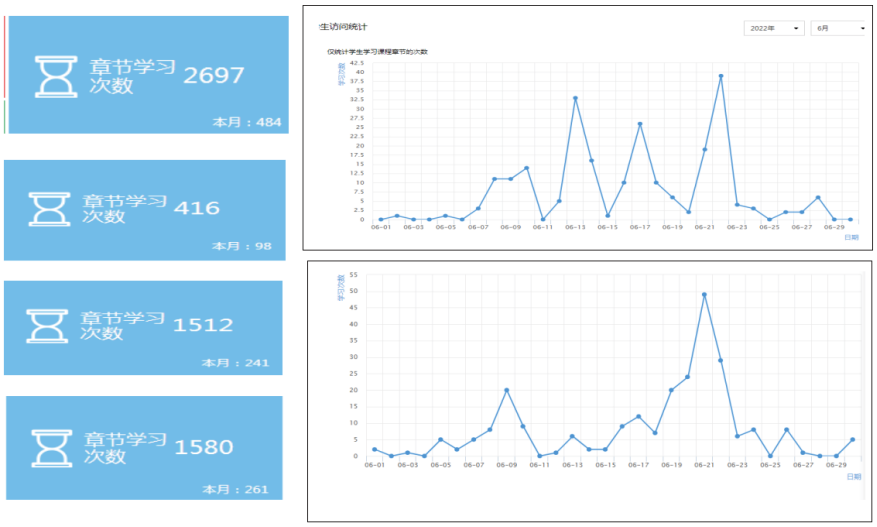

Figure 6 Statistics on Ss Using for Online Tests Banks in Xuexitong

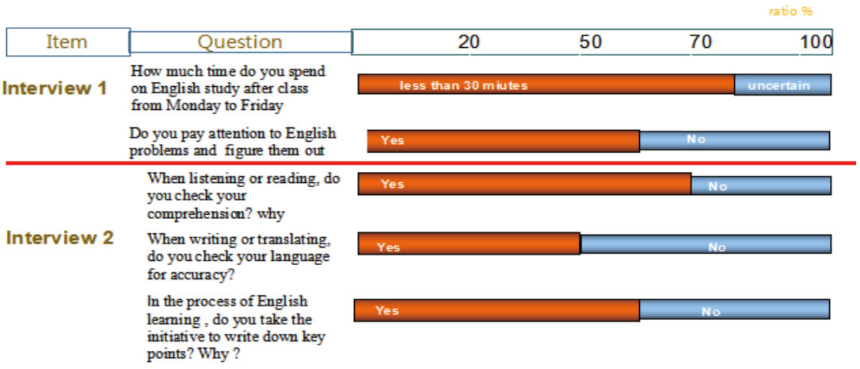

Figure 7 Interviews on independent Learning Initiative for Samples

Test application and analysis

From March 2022 to June 2022 (Semester II, grade 2021), the research group has conducted the diagnostic tests on sample classes while strictly implementing the teaching plan, and apply the results of assessment analysis directly to classroom teaching. The results of Unit 2, Unit 5 and Unit 8 are shown in Table 8, 9 and 10.

Figure 8 Average of Diagnostic Testing 2

Sample	Listening	Grammar	Translation	Total
ClassA	5. 32	30. 38	22. 84	58. 54
ClassB	6	31	27. 5	64. 47
ClassC	3. 85	29. 35	23. 9	57. 1

Table 9 Average of Diagnostic Testing 5

Sample	Listening	Grammar	Translation	Total
ClassA	5. 6	45. 62	19. 36	70. 58
ClassB	6. 6	50	22. 74	79. 34
ClassC	5. 95	41. 66	20. 93	68. 55

Table 10 Average of Diagnostic Testing 8

Sample	Listening	Grammar	Translation	Total
ClassA	4. 48	40	21. 92	66. 4
ClassB	6. 8	49. 11	30	85. 9
ClassC	4. 73	49. 55	20. 35	83. 63

According to the results in Table 8, 9 and 10, it is believed that the teaching arrangement for class A, B and C in Semester II is appropriate, and the diagnostic tests implemented are effective. Although the grammar and translation of Test 3 are greatly improved, and the listening is slightly improved, the effect is not as remarkable as grammar and translation.

Assessment Effectiveness and Acceptance

After the trial for one semester, the samples' final paper exam scores and final assessment scores both improved significantly. The worse the foundation of the samples, the greater the progress, and the better the foundation, the less progress (seen Table 11). The average score of the final paper exam of Class B increased from 31 points in the first semester to 49. 29 points, an increase of 48. 29%, and the assessment score increased from 55. 86 points to 68. 43 points, rising by 22. 50%. On the contrary, the final exam score and assessment score of Class C, which had a relatively good foundation, were both lower than those of the other two classes. From the cognitive perspective of view (seen Figure 12), students think that they have more progress than the first semester; the final assessment score is consistent with individual English competence; besides, most of them hold positive attitude to regular assessment.

Table 11 Comparison Between Average Final Exam and Assessment

Sample	Average Final paper Exam Score			Average Final Assessment Score		
	SemesterI	SemesterII	Ratio increased	SemesterI	SemesterII	Ratio increased
Class A	38. 30	46. 60	20. 36%	58. 70	68. 70	17. 04%
Class B	31. 00	49. 29	48. 29%	55. 86	68. 43	22. 50%
Class C	46. 63	53. 00	13. 66%	66. 13	71. 50	8. 12%

Concusion

By conducting the intelligent diagnostic tests, making use of intelligence technology to collect and analyze the samples, more detailed and personalized feedback information are provided by the system.. The research group traces the samples via the regular interviews and the analyzed data which are directly applied in the classroom teaching afterwards. Thus, it is can be found: 1) in a learning cycle and the “Lower students” have achieved a more remarkable improvement in grammar and translation,, on the contrary listening is less ob-

Vol.3 No.3 2022
https://doi.org/10.37420/j.cer.2022.028

vious; 2) The teacher's attention and external stimulus have a positive correlation with the learning internal cause for the intermediate and lower students, which will stimulate the students' subconscious to take action and make some changes in English learnign. The learning attitude and learning method of these samples change from passive to active. 3) Compared with traditional language testing methods, “online + pen-and-paper language testing” can better guarantee the implementation of "diagnosis - learning - research - teaching". At the same time, much more attention on the intermediate and lower students will have a negative the impact on the top students. In the further research of language testing, how to better collect language samples, evaluate the quality of language, interpret and report the evaluation results more effectively will be challenged by the advanced statistical assessment technology and information technology.

**This research is supported by the 2022 Scientific Research Fund of Education Department of Yunnan Province Study on The Assessment System of Vocational English Teaching Based on Intelligent Language Diagnostic Test (Project No. 2022J1674; Chairman: Zheng Xiao-yan).*

References

[1]. Bachman, L. F. (1990). *Fundamental Consideration in Language Testing*. Oxford: Oxford University Press.

[2]. Cambridge Assessment English. (2018). C2 Proficiency. http: // www. cambridgeenglish. org/exams-and-tests/proficiency/[2018-4 - 25].

[3]. Chapelle, C. A. (1998). Construct definition and validity inquiry in SLA research. In Bachman L F & Cohen A. D (eds). I*nterfaces Between Second Language Acquisition and Language Testing Research*. Cambridge: Cambridge University Press, 32-70.

[4]. Chen, L. M.,&Yang, H. (2021). *Theory and practice of English Testing*. Beijing: Foreign Language Teaching and Research Press. (in Chinese)

[5]. Chen, H. L. (2009). A Preliminary Study on the Validation Model of Validity Equivalence Between Pen-based Language Testing and Computer-based Language Testing. *Foreign Language,* (03): 73-80. (in Chinese)

[6]. Du, J. B. (1999). Diagnostic Testing in Foreign Llanguage Teaching. Foreign *Language Teaching and Research*, (04): 40-43. (in Chinese)

[7]. Du, J. B., & Gui, S. C. (2000). A Practical Study of Computerized Reading Diagnostic Tests. *Foreign Language Teaching and Research*, 32, (05): 345: 350. (in Chinese)

[8]. He, L. Z., & Li, H. (2011). The Main Research Paradigm of Language Testing and Its Development Trend. *Journal of Zhejiang University,* (01): 131-139. (in Chinese)

[9]. Huang, R. B. (2021). Language Assessment. *Shanghai*: *Shanghai Foreign Language Education Press*, 2021. (in Chinese)

[10]. Jin, Y. (2011). The development of Language Testing in Half A Century --A Summary of the 33rd Language Testing Research Symposium. *Foreign Language Testing and Teaching,* (04): 56-59. (in Chinese)

[11]. Jin, Y., & Sun, H. (2020). The Development of Language Testing Theory and Practice in China in The Past 40 Years: Retrospect and Prospect. *China Foreign Language*, 17, (04). 4-9(in Chinese)

[12]. Jin, Y., Wang, W., & Yang, H. Y. (2021). The Application of Technology in Language Testing:

Based on the Practice of CET-4 and CET-6. *Foreign Language Testing and Teaching*, (01): 2-7. (in Chinese)

[13]. Jin, Y.,& Zhang, L. (2016). The impact of test mode on the use of communication strategies in paired discussion. In Yu, G. & Jin, Y. (eds). Assessing Chinese Learners of English: Language Constructs, Consequences and Conundrums. *London: Palgrave Macmillan*, 61-84.

[14]. Jin, Y., Zhu, B., & Wang, W. (2017). *Writing to the Machine: Challenges Facing Automated Scoring in the College English Test in China.* Bogota, Columbia: The Language.

[15]. Li, X. J. (1997). *Language Tests and Art*. Changsha: Hunan Education Press. (in Chinese)

[16]. *Oxford Advanced Learner's Dictionary*. (2015). New 9th Edition. Oxford University Press.

[17]. Shohamy, E., Or, I. G. & May, S. (2017). *Language Testing and Assessment* (3rd ed). In May, S. (ed). Encyclopedia of Language and Education. Cham. Switzerland: Spring International Publishing. 149-161.

[18]. Zhu, Q.,&Shen Dan. (2020). Construction and Practice of O2O Intelligent Teaching Model of College English Listening and Speaking. *Teaching and Research*, (25): 232-234. (in Chinese)