Chinese 2003 Assessment

Overview

The goal of this assessment is to measure student learning received in the course of CHIN2003 - *Intermediate Chinese I* as a result of comparing students' performances in the pre- and post-tests.

The total number of the enrolled students in *CHIN2003* in fall 2018 is 26, among which 25 participated in the pretest while 20 participated in the post-test. This report only assesses the performances of the 19 students who participated in both tests.

Methods

The course of *CHIN2003* covers the first five chapters of the textbook *Integrated Chinese Level 1 Part 2*.The tests were designed based on the course content and the expected student learning outcomes and both tests were administered in the form of in-class written exams. The pretest was given to students at the beginning of the semester on August 29, 2018 and the post-test was conducted in the final week on December 5, 2018. Both tests were assembled to have the same items in the same sequence.

The tests consist of four sections. In the first section students are asked to choose the correct words or phrases to fill in the blanks in sentences. In the second section students need to respond to five questions according to their own experiences. The third section is reading comprehension. Students are required to answer true/false questions or to answer the questions in English (so they will not just copy sentences in the given paragraph) based on the two paragraphs they read. The fourth section is translation. It has two parts, Chinese-English translation and English-Chinese
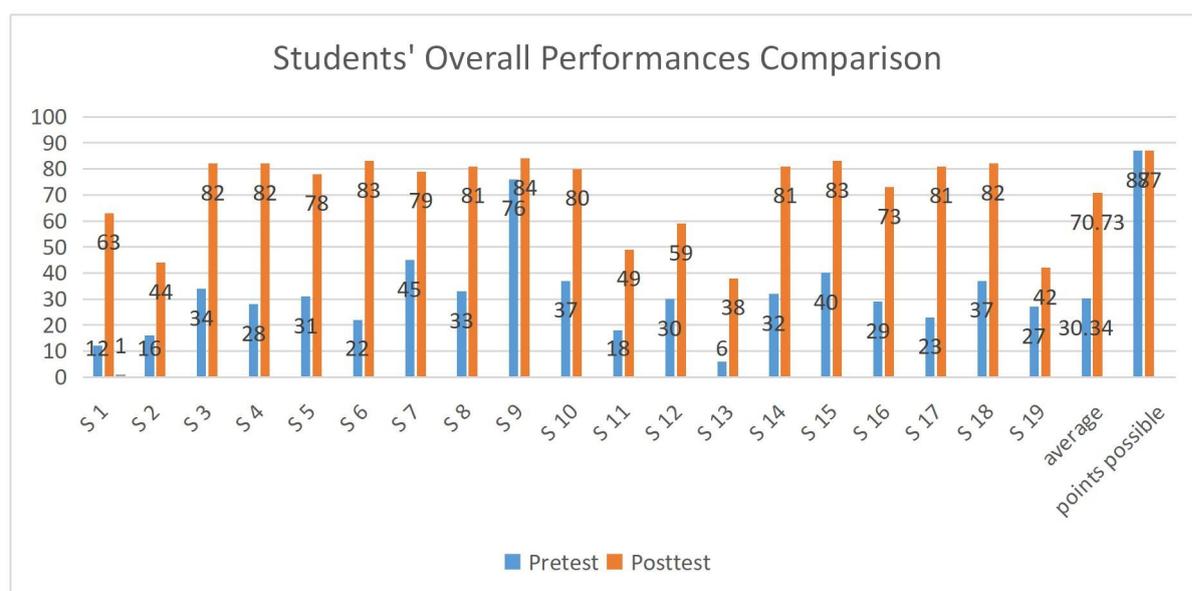
translation. The maximum possible points of both tests are 87.

In this assessment we first compared the overall results of both pre- and post-tests for each student to determine their improvement. We also conducted a comprehensive review across each section the post-test to see if students performed better in some sections than others. With this information, instructors could make alternation in their teaching to improve identified weaknesses or use the information about student strengths as the evidence of effective teaching.

Results

Students' Overall Performances

Figure 1. Comparison of Individual student overall performances in the Pre- and Post-tests

As Figure 1 shows above, all students performed better in the post-test. In the pretest, only one student (S 9) passed the test (the passing score is 87*60%=52.2, and S 9's score is 76). This student comes from Singapore and she grew up using Chinese, so it is not surprising that she scored as high as 75 in the pretest. In the post-test, 11 students scored more than 90% of the possible points (87*90%=78.3), 13 students
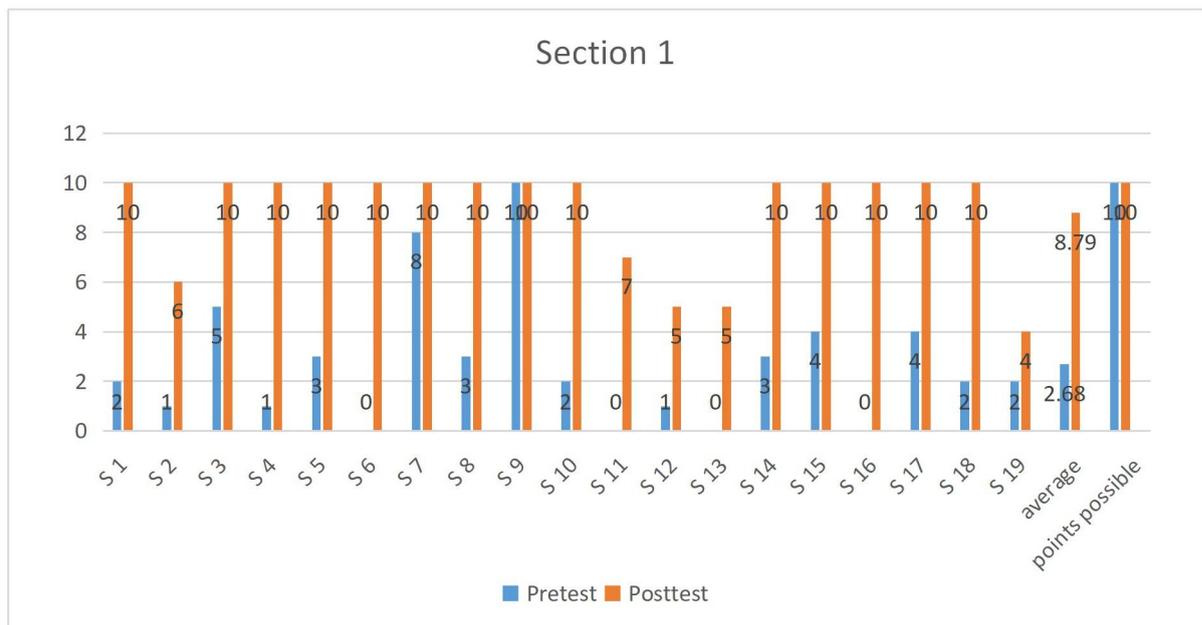
scored more than 80% of the possible points (87*80%=69.6), 14 students scored more than 70% of the possible points (87*70%=60.9), and 15 students scored more than 60% of the possible points (87*60%=52.5). Four students failed the post-test, but they all made some improvements.

The average points of the pretest are 30.3 and the average points of the post-test are 70.7, which shows an increasing rate of 133%. It's safe to say that students' reading and writing skills significantly improved through course learning and instruction.

Students' Performances in Section 1

Section 1 gives students ten words/phrases and asks them to choose the correct ones to fill in the blanks in sentences. This section is designed to test students' understanding of the key words and phrases they have learned in *CHIN2003* and of how those words and phrases work with other sentence elements.

Figure 2. Student Individual Performances on Section 1

As indicated above, all students did better in the post-test than in the pretest. There are14 students who received full points in this section and the average points are 8.79 out of 10.

Students' Performances in Section 2

Section 2 asks students five questions and students should answer those questions according to their own experiences. The objective of this section is to test students if they could understand questions with certain grammatical structures and if they could make appropriate responses. Students are required to write their responses in characters. Therefore, this section examines students' abilities of writing characters as well.

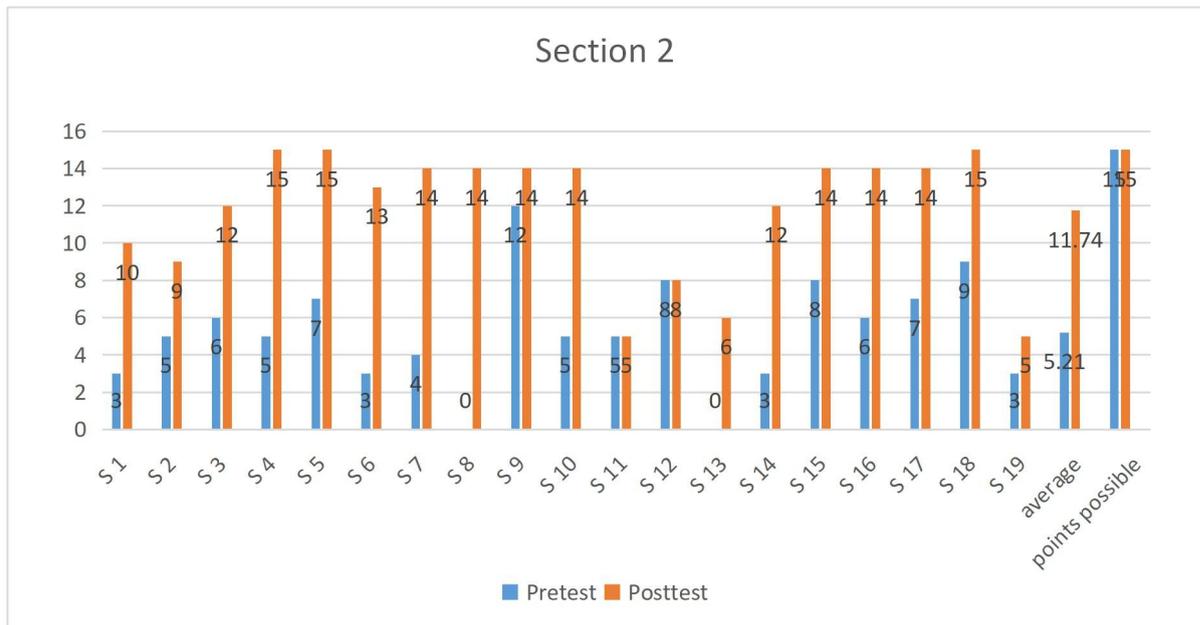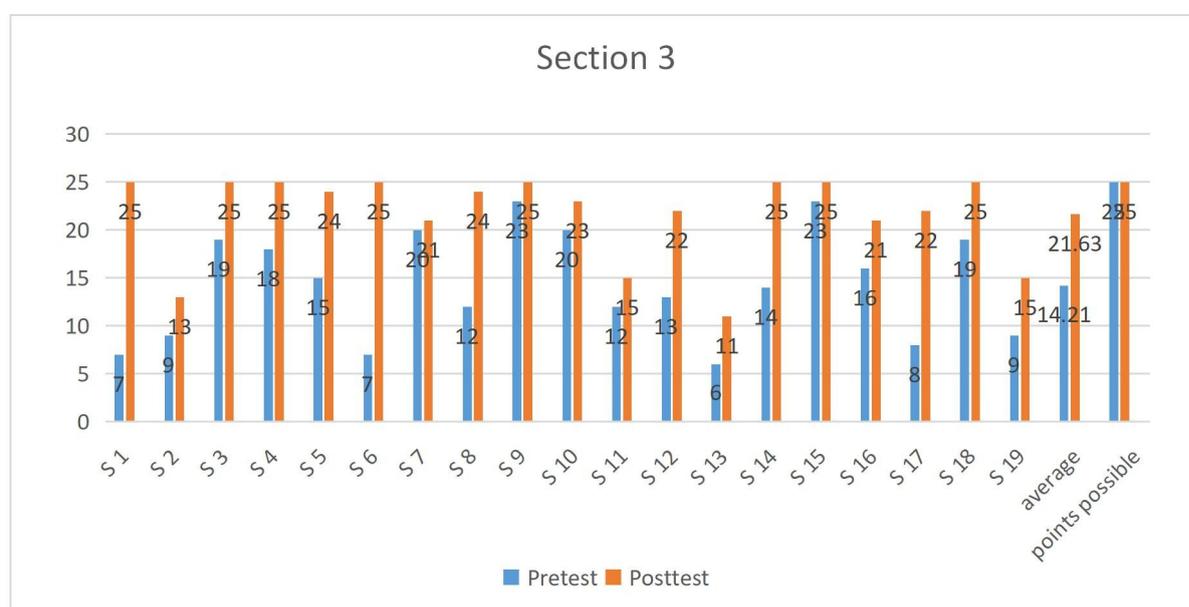Figure 3. Student Individual Performances on Section 2

Figure 3 shows students' performances on Section 2. All students scored higher in the post-test than in the pretest except S 11 and S 12. They received the same points in the pretest and the post-test. The average points students obtained on Section 2 are 11.74 out of 15, and only three students received full points of 15. One of the reasons that students did not perform as well on Section 2 as they did on Section 1 could be that in Section 2 students had to produce language output while in Section 1 students only needed to pick a given word or phrase to fill in a blank in a given sentence, which involved zero output. Another reason could be that students had to write characters in Section 2, which is well-known to be challenging to native English speakers.

Students' Performances on Section 3

Section 3 is reading comprehension, which only tests students' reading proficiency. Students' language output is not required in this section. Students will read two paragraphs. They need to answer five true/false questions to five statements based on the first paragraph and they are going to answer five questions after they read Paragraph 2. The questions are in Chinese but students will answer them in English.

Figure 4. Student Individual Performances on Section 3

Based on the information presented in Figure 4, one can see that students showed better performances on Section 3 than on Section 2. The average points in the post-test are 21.63 out of 25, which make the accuracy of 86.4%, while the average accuracy on Section 1 and 2 in the post-test is 88% and 78%, respectively.

All students made improvements in the post-test. However, the differences between the pretest and the post-test on Section 3 were not as obvious as those on other sections. Because in the pretest, students did noticeably better on Section 3 than on other sections. The average accuracy on Section 3 in the pretest is as high as 56.8% (14.21 out of 25), while this data is 34% in Section 2 and 26.8% in Section 1.

The fact that students did better on Section 3 in the pretest than on the other sections
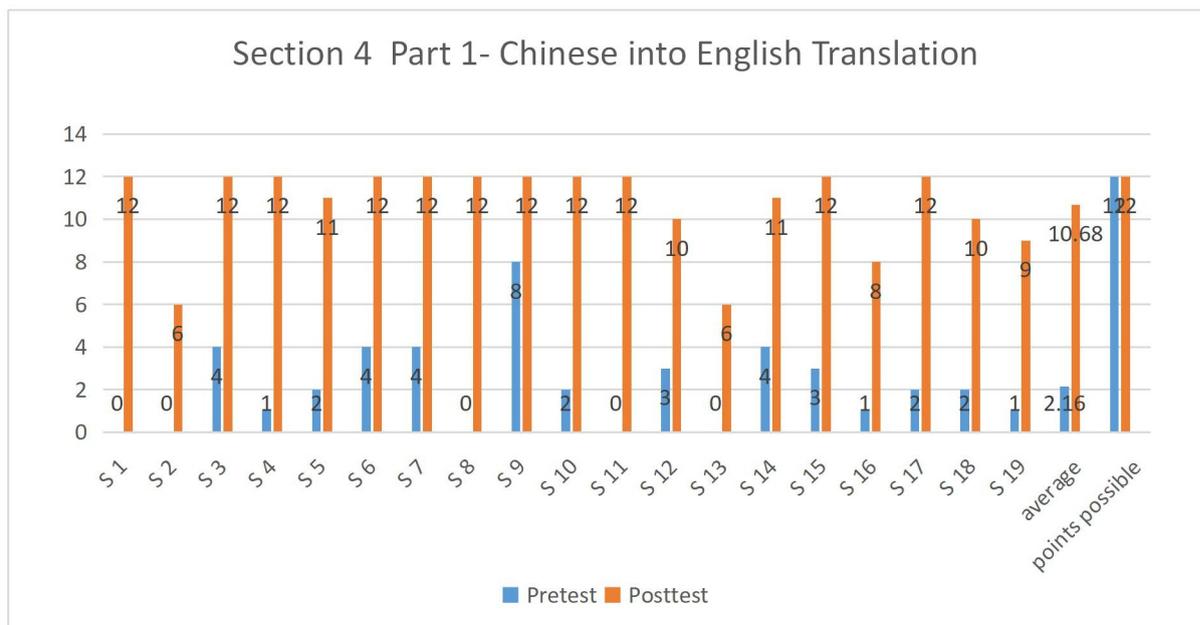
could be explained as the following. In Section 3, students are provided with paragraphs instead of unrelated single sentences. Therefore, they will inevitably encounter many words and phrases they learned in Elementary Chinese, which reveals to students a certain amount of information that they can rely on to sort out the general idea of the passage.

In the post-test, however, students did not do as well on Section 3 as on Section 1. This result was expected since reading passages in Chinese characters could be intimidating and overwhelming, especially during an exam. It is easy to overlook some details when students have to deal with a great amount of information.

Students' Performances on Section 4 Part 1- Chinese into English Translation

In this part, students were given three Chinese sentences with some grammar structures and vocabulary they will learn or learned in *CHIN2003*. It looks like another reading comprehension section, but sentence translation examines students' understanding of certain language elements more specifically.

Figure 5. Student Individual Performances on Section 4 Part 1 - Chinese into English Translation

As Figure 5 shows, the average points students received on Section 4 Part 1 were 10.68 out of 12, which makes the accuracy 89%, the highest across all sections. One
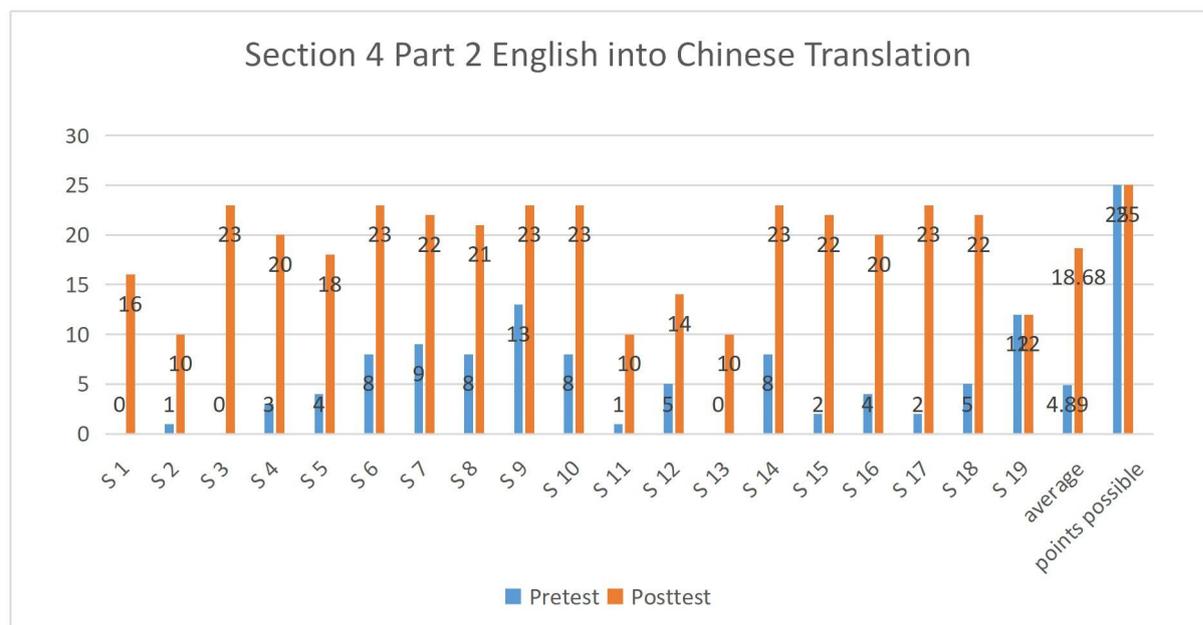
can also see the greatest improvement among all the sections since the average accuracy on Section 4 Part 1 in the pretest is only 18% (2.16 out of 12). There are 5 students who did not get any point in the pretest, but in the post-test, 3 of them received full points of 12.

As discussed above, this part was designed to aim specifically at vocabulary and grammar structures students will learn or have learned in *CHIN2003*. The result in the pretest was fully expected and the result in the post-test could be used as the evidence of successful acquisition and effective teaching.

Students' Performances on Section 4 Part 2- English into Chinese Translation

In this part, students are required to translate five English sentences into Chinese and they need to hand-write their answers in characters. This part examines students' grasp of the usage of the vocabulary and grammar structures taught in CHIN2003, and at the same time, tests students' character writing skills.

Figure 6. Student Individual Performances on Section 4 Part 2 English into Chinese Translation

As shown above, all students made improvements in the post-test on Section 4 Part 2 except S 19 who got 12 points in both tests. The average points students scored on this part are 18.68 in the post-test, and the average accuracy is 74.7%, the lowest

among all sections. No one received full points. The average accuracy in the pretest is 19.6% (4.89 points out of 25), a little higher than the accuracy on Section 4 Part 1 in the pretest.

This part requires students to produce output based on the given English sentences. It has been proven to be the most difficult part of the test. There are two reasons why this part is more difficult than Section 2. The first one is that the questions in Section 2 are of great assistance in students' output. Because in Chinese, the sentence order does not change in questions, students only need to identify the question pronouns and simply replace them when answering questions. The second reason lies in the fact that students enjoy a certain degree of freedom to pick the words they know when answering questions in Section 2. For example, to answer the question "Where do you often put the fruits you buy?", if a student does not remember the word for "refrigerator", he/she can change the answer from "in the refrigerator" into "on the table". However, in Section 4 Part 2, students will have to translate the given sentences, which leaves them little space to bypass the words or structures they do not know or remember. The common situation is that they either know how to translate a sentence, or they don't.

Final Thoughts

Both students' overall performances and their performances in each section in the two tests have shown significant improvements in their Chinese reading and writing proficiency. To give a holistic assessment of learning outcomes and teaching effectiveness, the future pre- and post-tests should also incorporate sections that examine listening and speaking skills.

Although every student's points increased in the post-test, the fact that 5 students received lesson than 70% of the points should not be neglected. The Chinese program has been trying to offer additional help to students by organizing Chinese Lab,

Chinese Conversation Table and Peer Tutoring sessions, etc. Instructors can encourage those students to take advantage of those resources.

The last thought concerns Chinese characters teaching. Among those 5 students who received less than 70% of the points, 2 of them (S 13 and S 19) are heritage students who have comparatively high speaking proficiency. It's obvious that characters are the obstacles for them to reach a higher level of reading and writing. Unfortunately, their characters did not improve much after one semester's learning while other heritage students (S 6, S 10, S 15, and S 16) demonstrated a more advanced level of literacy at the end of the semester. It is definitely worth thinking about how to motivate students and raise their interest in character learning.

French 2003 Assessment

Overview:

The Assessment Exam, intended to gather data for the study, was given twice during the semester: the first time during the second week of class, on January 23, 2019, and the second during the last full week of class, on April 22, 2019. Thus students were tested approximately three months apart, at the beginning and the end of the course. Students were apprised of the nature of the exam and its purpose, and were requested to take the exam seriously, although it was not for a grade. To maximize participation, a small percentage of their final grade was based on simple completion of both exams; for the same reason, the exam was given in class, and had every appearance of a typical exam in the course. It had been suggested that an exam on Blackboard might accomplish our goals, but the test planners adjudged that students might not take seriously an exam in such a format. Our belief was that if it had every appearance of a normal exam, students might unconsciously take it more seriously, and give their best effort.

Assessment Instrument:

The Assessment Instrument for the exam was based upon the final exam of the preceding course (French 1013 – Elementary French II). The test planners made an effort to eliminate subjectivity in evaluating students' performance, and therefore eliminated the writing or "composition" section from the exam. Most other elements of the French 1013 final were retained: reading section, listening comprehension section, and lengthy grammar section.

To create an Assessment Instrument that could be completed in a 50-minute period, as well as one that could be objectively graded, the planners changed the format of student answers. On the French 1013 final, grammar is tested by a variety of means, but all require student input: fill in the blank of verb forms or vocabulary, rewriting of sentences using negation or pronouns, question construction, and so forth. For the Assessment Instrument, the planners converted the grammar sections to "recognition-only," using a multiple-choice format. Students did not have to produce correct grammar forms; rather they identified correct forms in a pool of four answers. Only one answer was correct for each item, with a variety of incorrect or 'distractor' answers. Although this approach might be considered a simplified one, it enabled objective grading, and since it was employed on both exams, the students' task was the same each time. The planners expected this to provide an accurate baseline for the results.

In the interest of objective grading, the planners provided Scantron forms to the students. The experts at Gibson Annex ran the forms for our Department, providing us with the data that follows.

Results of First Assessment Exam:

The number of students tested and the mean percentage results are as follows:

Section:	Enrollment:	Mean Percentage per section:
001	9	46.22 %

002	13	62.84 %
003	13	51.92 %
004	9	50.56 %
005	13	52.92 %

Results of Second Assessment Exam:

The number of students tested remained the same. The mean percentage results are as follows:

Section:	Enrollment:	Mean Percentage per section:
001	9	52.25 %
002	13	71.36 %
003	13	57.64 %
004	9	53.78 %
005	13	60.15 %

Comparison of Results of the two Assessment Exams:

The results side-by-side and the variance are as follows:

Section:	Mean, Exam I:	Mean, Exam II:	Variance:
001	46.22 %	52.25 %	+ 6.03 %
002	62.84 %	71.36 %	+ 8.52 %
003	51.92 %	57.64 %	+ 5.72 %
004	50.56 %	53.78 %	+ 3.22 %
005	52.92 %	60.15 %	+ 7.23%

Thus, the mean variance across the five sections was + 6.14 %. Students scored over six percentage points better on the same exam after taking the course.

Conclusion:

This variance is positive, and we see improvement across all sections. The clear conclusion is that students perform better at recognition of basic structures and vocabulary in French after completing French 2003.

That we do not find greater improvement than an average of just over six percent could be attributed to several factors. In January, many students who enrolled in French 2003 had little more than a month previously taken the French 1013 final; perhaps it was still somewhat familiar, creating falsely high first results. Another possible factor is that while some of the basic structures students learn in French 1013 are specifically reviewed in French 2003, not all are; moreover, new concepts and structures are introduced. Clearly, what is new in French 2003 could not be tested before the first assessment exam; doing so would create a falsely low first result. The planners decided that using a modified French final 1013 exam would set a baseline for determining what students knew upon enrolling in the course. Of course, the two assessment exams had to be the same. The result of this necessity is that the second assessment exam tests material that may not be explicitly taught in French 2003, although it tests material that is essential to learning French.

The planners hope to develop improvements to the testing format before the next Bi-Annual Assessment, in order to diminish the factors mentioned above. It may be that the study of practices of other departments in the University or language departments in other institutions will enable us to improve our methods in the future. We welcome any comments or suggestions.

German 2003 Assessment

        Since Spring 2018, we have conducted pre- and post-assessments in GERM 2003, Intermediate German I, in accordance with the university's recent requirement that university core courses be assessed to measure student progress. To this end, the faculty in the German section created a two-part examination designed to assess students' knowledge of grammar, vocabulary, and culture in the first multiple choice "written" section, and students' listening and speaking abilities in a second recorded interview. The same assessment was given to students at the beginning and end of the semester to more accurately reflect change across the duration of the course. We have administered the assessment twice (Spring and Fall 2018), but are still working on refining the testing process (due to issues with the software, for example, several students had difficulty recording their oral interviews, thus leaving incomplete data from pre- to post-assessment; additionally, the assessment is not mandatory, and as such has inconsistent participation). Summary of findings: the data suggest that students are making improvement in both written and oral skills from the beginning of GERM 2003 to the end, though there is room for greater improvement. The assessment has not yet yielded sufficient data, however, to identify trends in student performance. In order to make concrete changes to the curriculum at this level, it will be useful to develop a more precise assessment module and rectify the issues preventing some students from completing both pre- and post-assessments in either written or oral proficiency. Below are the results from the last two assessment periods:

Spring 2018

Pre-Assessment

        28/36 passed written (78%)

        30/30 passed oral (100%)

The average score for writing was 68%; the average score for speaking was 83%.

Post-Assessment

        32/34 passed written (94%)

        22/22 passed oral (100%)

The average score for writing was 76%; the average score for speaking was 84%. From pre- to post-assessment, students improved an average of 8% in writing, and 1% in listening, while the passing rate in writing increased by 16%. The minimal improvement in oral proficiency may be due to a sharp drop in completion of the oral component on the post-assessment.

Fall 2018

Pre-Assessment

        37/47 passed written (79%)

        20/20 passed oral (100%)

The average score for writing was 66%; the average score for speaking was 95%.

Post-Assessment

    24/31 passed written (77%)

    22/24 passed oral (92%)

The average score for writing was 69%; the average score for speaking was 91%. From pre- to post-assessment, students improved 3% in writing, but dropped 4% in listening. The latter value is due in part to an increase in participation on the oral component during the post-assessment, which introduced students with weaker scores than those who had completed the oral pre-assessment. The passing rate for both modules dropped from pre- to post-assessment, but we have yet to identify specific areas for increased focus in the written assessment.

Italian 2003 Assessment

Italian 2003 Assessment, Fall 2018

Dr. Rozier and Mr. Anthony Sargenti gave a short True/ False, Multiple Choice and Writing pre-assessment test on August 24th; the post-assessment test was given on December 3rd. There were 2 sections of ITAL 2003, with **30 students.** Students completed exercises to evaluate their skills in listening comprehension, vocabulary, grammar, reading and writing (T/F and short answers), and culture.

	Pre-Assessment Score	**Post-Assessment Score**	**Improvement**
Listening	51%	99%	48%
Vocabulary	46%	86%	41%
Grammar	31%	56%	25%
Reading/Writing	59%	86%	26%
Culture	56%	64%	8%
Total	42%	67%	25%

Improvements were made in all categories. Going forward we will change the cultural exercise from T/F to multiple choice or short answer, since there could be false positives in the pre-Assessment.

A copy of the assessment is included.