Language Testing: An Example Placement Test

Matthew Miller

Introduction

Baker (1989: 5) states that "the surefire way of placing a learner in a school... would be to put him in a class and see how he gets on, moving him if necessary." He admits that this is "time-consuming and inconvenient" and that is why placement tests exist. Placement tests are used for their "ability to spread students out according to ability on a proficiency range within the desired area of learning" (Henning, 1987: 6). This essay proposes a two-part placement test for an English conversation course in Japan by first explaining the test's purpose and describing the students and the course. Then the test is presented, tasks justified, and in conclusion, ways to investigate its reliability and validity are discussed.

1 . Test purpose and student/course description

The purpose of the presented test is to place new and returning students into the appropriate levels of an English conversation course according to grammatical, listening and speaking ability. Brown (1996) said that ideally the placement tests should be written and given before the development of the class materials to ensure that the materials are able to achieve the objectives. However, this test was written for a course already in existence whose materials were previously chosen. The placement test this language program employed was merely a ten question, grammar-based, multiple-choice written assessment scored by the school staff who then used the results to decide student placement. However, instructors recurrently complained that students were often misplaced in a level that did not reflect their speaking ability. This inspired the investigation into a more reliable and valid process of placement.

The proposed two-part placement test is designed to indicate the examinee's approximate level of grammar knowledge, comprehension and speaking ability. The scoring system of Part A will tentatively place the

students into a level where the grammar covered in the corresponding text would be only review and the listening activities fairly untroublesome. Part B examines how well the students are able to orally express themselves in regards to different situations and language. This material was taken from speaking activities in the text. It will provisionally place them at a level in which they were not able to, on average, satisfactorily respond to that level's material. The final placement is then decided (see 3.3.4 *Final Placement*) hopefully settling the students into a place where they are already proficient in the grammar but wherein the speaking activities remain challenging.

This course, part of a larger group of English language programs called 'Senka Intensive Studies', offered a year-long curriculum at a chain of private language schools across Japan. It was divided into 10-lesson trimesters and targeted towards intermediate-level students wanting to improve their speaking ability. Class size was limited to 12 students and although they could enroll at any time, the staff would recommend the new student join on the first day of a trimester. The class met for three hours (with two 15-minute breaks) for ten consecutive Saturdays and had the same native English-speaking teacher for at least one trimester. The program consisted of three levels: Level 1 pre-intermediate, Level 2 intermediate and Level 3 upper-intermediate. Beginning and advanced level students were encouraged to join different programs. The texts used were as follows:

Level 1	*New Interchange 2* (Cambridge University Press)
Level 2	*New Interchange 3* (Cambridge University Press)
Level 3	*Passages 1* (Cambridge University Press)

The course objective was to improve intermediate-level students' conversation skills through brief reviews of previously studied grammar, some listening exercises and many varied speaking activities. The texts for each level contained grammar that the students should have already sufficiently comprehended, thus allowing most of the class time to concentrate on speaking skills.

At the beginning of each trimester, the students had a one-hour orientation given by staff and teachers in English and Japanese introducing them to the class rules, teachers' expectations and course objectives. Prior to this, the course's instructors were given a five-hour training by the program

coordinator who was also a native English speaker that involved practicing group teaching skills, student counseling procedures, course and text-specific information, and included supplementary speaking activities. Training in administering placement tests also needed to be introduced due to the significance of test interpretability (Henning, 1987).

The students were almost all Japanese ranging in age from 18 to about 60, with the average age being 25-30. A large majority of the students had taken classes at this particular language school before and most had some university education and experience traveling overseas. They joined the program for business or educational reasons yet did not have time to enroll in one of the regular intensive courses that met five times a week.

Generally, Japanese students study English grammar and vocabulary from elementary school through university. These classes are rarely communicative resulting in students with higher skills in grammar and vocabulary than in speaking. Thus the popularity and importance of conversation courses similar to this one that can assist students in using the knowledge they already have and turning it into a tool for conversation and communication.

2 . The placement test
2 . 1 Administration
The proposed placement test for the course consists of two distinct parts (see Appendices A and B). Part A is administered and graded by school staff. It is a multiple-choice test of grammar knowledge and listening skills and can be taken at any time during school hours in a private classroom or listening booth. The staff explains the test directions in Japanese before starting the listening cassette. Part B is taken after the first has been graded and a tentative level assigned. It is administered orally by one of the native English-speaking teachers to assess speaking ability. This test must be made by appointment at least one day in advance to fit into the teacher's schedule. The instructor, considering both test results, then decides where to place the student and informs the staff. Any discrepancies, such as limits in class size, can be resolved through discussion between the Senka program coordinator, the staff and the teacher who administered the test with the school director making the final decision.

2 . 2 Part A (Appendix A)

Since the focus of the program was to improve speaking skills, the students must already posses some proficiency with the grammar and listening exercises presented in the text. This first test is designed to grade both at the same time. The listenings (questions 1 - 6) are an *integrative test* measuring several skills at once: grammar, vocabulary, and comprehension. While the second part (questions 7 -15) is a *discrete-point* test that determines the students' skills in a limited range of language. In this case, the test covers only grammar taken directly from the texts (Henning, 1987).

The conversations are heard once with five second intervals for the students to answer. They are then repeated without the intervals to allow the students to check their answers. The students are given 15 minutes to complete the test. This places it closer to a power test in which the students are allowed to finish, while containing several questions that most students are not expected to be able to answer correctly, than a *speed test* which contains questions most students can answer correctly but perhaps do not have enough time to finish (Henning, 1987). The students are not told the result of Part A until Part B is completed.

The questions are multiple-choice for the reasons Oller (1979) recommended; they are objective, reliable in scoring, and easy to administer and grade. Because these tests were handled by school staff who may not speak very much English, the ease of administration and grading the test are especially important.

Baker (1989) believes an ideal placement test will discover how familiar the students are with the course material, thus all of the conversations and grammar questions are taken directly from the first half of each of the three texts. For some of the listening questions, the students must be able to understand the listening *and* grammar to answer correctly. For example, to answer the first question correctly, the students must have heard "used to" as well as understand that it describes an action in the past that is no longer performed. Each question is based on separate material so that if one question is not understood, the students still have a fair chance of getting the next question correct and also to test the students' familiarity with a broader range of language. All of the answers are straightforward avoiding any sense that there could possibly be more than one answer. These factors help to

increase the test's reliability (Hughes, 1989).

A score of 0 - 5 on Part A places a student in Level 0 (meaning a beginners' course may be more suitable) until the second test is given. A score of 6 or 7 places him or her in Level 1 , 8 -12 in Level 2 , and 13-15 in Level 3 + (Level 3 or a higher-level program depending on the results of the oral test). Thus, Part A of the test gives some idea of the student's listening ability and how familiar the student is with grammar presented at each level of the course. This scoring system then places the student in a level where theoretically he or her is familiar with at least half of the grammar.

2 . 3 Part B (Appendix B)

The aim of this test is to gauge the students' speaking ability through a series of questions based on subjects and language used in the three texts. Harrison (1983) considers getting the students to understand what is expected of them an obstacle to communicative testing. To avoid this, the staff explains the procedures in Japanese prior to the test and the instructors are trained to model examples effectively if the students are still unsure of what to do. The native English-speaking instructor who gives the test will not have seen the result of the students' first test to prevent "possible scorer bias" (Harris, 1969: 92).

The instructor and students are in a booth or private classroom at a table sitting across from each other. The warm-up questions relax the students and get them thinking in English (Weir, 1993) before posing the Level 1 questions. The teacher should keep the score sheet on its clipboard and in his or her lap so the students cannot see it.

The grading system was based on Harris' (1969: 84) 'Sample Oral-English Rating Sheet' and is explained in the teacher's instructions at the top of the test (see Appendix B). The question or statement is read and repeated if so requested by the students. The instructor is trained to react as if in a natural conversation perhaps agreeing, sharing his or her own opinion, giving a personal example, etc. for up to half a minute before assigning a grade and moving on. This test should last 5 -20 minutes. Keeping in mind the students' ability to comprehend what was asked of them, fluency, ability to expand, pronunciation and vocabulary, the instructor decides the score and

records it. Zero points are given if an inappropriate, incomprehensible or no reaction was given. One point is awarded to a satisfactory response and two points if the students answer or explain themselves very well.

If a student receives a total score of two points or less in Part B's Level 1 section, the test is finished and the staff is asked to recommend a lower-level program regardless of Part A's outcome. A score of three to nine also ends the test and it is suggested that the student is assigned to Level 1 . Once a score of ten is reached, even if all ten questions have not been asked, the instructor is to move on to the Level 2 questions. A score of 0 - 9 in the next set of questions suggests the student be placed in Level 2 and once 10 is reached, the instructor should begin asking questions from the Level 3 section. A score of 0 - 9 suggests the student should be placed into Level 3 . A score of 10 or more will end the test and the student will be recommended to join an advanced level program regardless of Part A's result.

2 . 4 Final Placement (Appendix C)

After a suggested level has been assigned from Part B, the instructor then checks Part A's suggested level and uses the Final Level Placement Chart (Appendix C). This chart finds the correct placement by considering the student's grammatical knowledge, comparing it with speaking ability, and assigning an appropriate level where the language focus is likely a review but the speaking exercises are demanding.   Once the trimester has started and a teacher feels that one or more students have been misplaced, after 2 lessons (six hours of class time)a level-charge request may be submitted to the coordinator. The coordinator will then alert the appropriate school director who then considers class size and discusses the matter with the student.

Thus, for example, someone scoring very low on Part A and very high on Part B, which is an extreme and unlikely combination, would be placed in Level 2 . Then, if that level was unsuitable, the student would be able to move up or down after two weeks. This procedure is a safety valve for cases of students who do not test well, teachers not grading appropriately, or any of the numerous reasons for inaccurate results discussed in the following section.

3 . Reliability and validity

> "The most important quality to consider in the development, interpretation, and use of language tests is validity.... While validity is the most important quality of test use, reliability is a necessary condition for validity."
>
> (Bachman, 1990: 289)

3 . 1 Reliability

Bachman (1990: 240) says that reliability is "agreement between similar measures of the same trait" and that validity is "agreement between different measures of the same trait". Factors that can damage test reliability include the students (i.e. they may be unwell on that particular day or familiar/unfamiliar with taking tests), scoring (staff making errors in grading Part 1 and the instructors subjectively scoring in Part B), and differences in test administration (tests given in different schools on different days and administered by different people would vary in conditions and errors). The small number of items in Part A and the 33% chance of guessing a correct answer also contribute to the unreliability (Henning, 1987). The following paragraphs contain ideas to investigate the proposed placement test's reliability and validity.

A way to check reliability is the test-retest method, where "the same test is readministered to the same people following an interval of no more than two weeks" (Henning, 1987: 81). Due to logistical difficulties, doing this would be impractical in this particular course's situation. However, the following is an adaptation of the test-retest method that would give similar results.

New Interchange 1 and 2 provide tests for every four units of text. The first test is administered on the first day of class (a similar test is written and given for Passages). The students are placed into pairs and asked to perform in front of the teacher a speaking task based the same ten questions/subjects that were asked for the corresponding level in the placement test. The instructor evaluates the students based on the same stipulations stated in Part B of the placement test. The results are recorded by the instructor and sent to the Senka coordinator who would then use Henning's (1987) formula and all of the schools' information to find out if the placement test was reliable.

In addition, at the end of each trimester a survey is given to each student

and instructor. It includes the following questions: 'Were you/the students placed in the proper level?' and 'If not, in what level do you think you/the students should have been placed?' The Senka coordinator compares these surveys with the results of the corresponding placement tests and the first day tests then attempts to conclude where the problem is and how to revise the placement procedure effectively.

3 . 2 Validity

"A test is said to be valid to the extent that it measures what it is supposed to measure" (Henning 1987: 89). The test content itself is not being measured for validity, but the way in which the tests results are "interpreted and used" is what is measured (Bachman, 1990: 238). There are four items that are considered here when searching for validity; *content validity, criterion-related validity, construct validity*, and *face validity* (Hughes, 1989).

Having content validity, the placement test's "content constitutes a representative sample of language skills, structures, etc." (Hughes, 1989: 22). In order to check this, a chart specifying the exact skills and language structures should be made and compared to what is present in the test. These should directly correlate for it to be valid (ibid).

To test for criterion related validity, one investigates whether the test functioned properly. This is done in the form of the previously mentioned questionnaire sent out to all of the instructors inquiring how well the students were placed (Hughes, 1989). This question asks the teachers to list specific problems, i.e. grammar, listening and speaking, in the suspected misplaced students.

Finally, face validity, or how valid a test actually looks to the observer is very important. This can be resolved in a meeting between the instructors, staff and the coordinator where the new placement test is presented and explained. The attendees are given a chance to comment on and/or criticize the test allowing the coordinator to perceive the face validity (Hughes 1989).

Conclusion

Due to the numerous factors involved, designing effective placement tests is a challenging task. Finding the balance between a test that is reliable and

valid and realistically administrable requires a great amount of time and trial and error. However, this proposed placement test as outlined in this paper is an example of developing an improved system and allowing for future adaptation by nature of its flexibility and investigative proposals.

Bibliography

Bachman, L. (1990) *Fundamental Considerations in language Testing*. Oxford: Oxford University Press.

Baker, D. (1989) *Language Testing: a Critical Survey and Practical Guide*. London:Edward Arnold.

Brown, J. D. (1996) *Testing in Language Programmes*. Upper Saddle River, N.J.:Prentice Hall Regents.

Harris, D. P. (1969) *Testing English as a second language*. New York: McGraw-Hill.

Harrison, A. (1983) "Communicative testing: jam tomorrow?" Keynote paper. In Hughes and Porter (eds.) (1983) *Current Developments in Language Testing*.London: Academic Press (p. 77-85).

Henning, G. (1987) *A guide to language testing: development, evaluation, research* Boston: Heinle and Heinle Publishers.

Hughes, A. (1989) *Testing for Language Teachers*. Cambridge: Cambridge University Press.

Oller, J. W. (1979) "Language tests at school: a pragmatic approach" In C.N. Candlin(ed.) (1979) *Applied linguistics and language study*. London: Longman.

Richards, J., Hull, J. and Proctor, S. (1998a) *New Interchange 2* . Cambridge: Cambridge University Press.

Richards, J., Hull, J. and Proctor, S. (1998b) *New Interchange 3* . Cambridge: Cambridge University Press.

Richards, J. and Sandy, C. (1998) *New Interchange 2* . Cambridge: Cambridge University Press.

Weir, C. J. (1993) *Understanding and Developing Language Tests*. New York: Prentice Hall.

Appendices

Appendex A: Part A

Saturday Senka Placement Part A	
Listen to the conversations and circle the best answer	
1 . Conversation 1 Circle the best sentence. a. He spends two weeks at his uncle's beach house every summer. b. He doesn't spend two weeks at his uncle's beach house anymore. c. He is going to spend two weeks at his uncle's beach house.	**9 . Choose the best answer:** The Mona Lisa ____ by Leonardo da Vinci. a. has painted b. painting c. was painted
2 . Conversation 2 Choose the best sentence. a. There are too many busses and cars. b. There is too little traffic and too few cars. c. There aren't enough subway lines.	**10. Choose the best answer:** I ____ with my friend yesterday, and I ____ my keys. a. was shopping dots ... lost b. shopped ... lose c. shop ... had lost
3 . Conversation 3 She doesn't like guys ... a. she has something in common with. b. who are too serious. c. who she can talk to easily.	**11. Choose the best answer:** The oceans ____ by industrial waste. a. pollute b. are being polluted c. are polluting
4 . Conversation 4 Choose the best answer. An intern ... than a landscaper. a. has better hours b. is better paid c. is harder	**12. Choose the best answer:** Joe: Would you rather ____ English in Canada or Australia? Mary: I'd prefer ____ in Australia because it's warmer. a. learning ... study b. to learn ... to study c. learn ... studying
5 . Conversation 5 What did Carol learn? a. That opposites attract. b. To appreciate her parents and friends. c. How to get along with others	**13. Choose the best answer:** ____ the city is safe, you should still be careful. a. Even though ... b. Because ... c. However ...

6 . Conversation 6

What is different about their families?

a. She sees her family a lot.

b. His family is small.

c. She isn't close to her family.

14. Choose the best answer:

____ I've had a cup of coffee in the morning, I am sleepy.

a. While…..

b. Until…..

c. As soon as….

7 . Choose the best answer:

Have you ever ____ sushi?

a. try

b. tried

c. trying

15. Choose the best answer:

I ____ more time studying in high school.

a. better to spend

b. ought to spent

c. should have spent

8 . Choose the best answer:

Sometime in the future____

a. I will travel to Europe.

b. I have traveled to Europe.

c. I travel to Europe.

SCORE: ____ / 15

SUGGESTED LEVEL ____

Scoring key (for staff)	
Score	Suggested Level
0 - 5	Level 0
6 - 7	Level 1
8 -12	Level 2
13-15	Level 3 or higher

Appendex B: Part B

Saturday Senka Placement Part B

*For each question assign the following points according to the student's ability. A 10/20 score or higher is a pass. Once the passing mark is reached, skip to the next level's questions. The highest level the student fails is the correct level for placement. If the student receives **less than** a 3/20 on the Level 1 questions, ask the staff to recommend a beginning-level Senka program. If 10 points is reached in the Level 3 section, ask the staff to recommend an advanced-level Senka program.*

0 points	Unable to understand the question, asked for the question to be repeated more than once, responded incorrectly or unnaturally, hesitated an abnormal amount of time, no expansion or follow-up question from student, poor pronunciation impeded communication, limited vocabulary and/or misuse impeded communication.
1 point	Able to understand the question, asked for the question to be repeated only once, the response was correct and fairly natural, hesitated somewhat in responding, no expansion or follow-up question from student, pronunciation did not impede communication, limited vocabulary and/or misuse present but able to get point across.
2 points	Able to understand the question, did not ask for the question to be repeated, the response was correct, very natural and fluent, did not hesitate in responding, expansion and/or follow-up question from student, pronunciation did not impede communication, no problems communicating due to vocabulary.

Warm-up Questions

1 . Hello, how are you?
2 . My name is _Steve_ . What's your name?
3 . I live in _Yokohama_ . Where do you live?
4 . I'm an English teacher. What do you do?

Level 1 Questions

#	Questions	Points
1 .	What did you used to do for fun as a high school student?	
2 .	What is one problem this city has and what are some ways to solve it?	
3 .	Would you prefer to live in a house or apartment? Why?	
4 .	How do you make your favorite snack or dish?	
5 .	What are you going to do next week?	
6 .	If you need to apologize for something, what do you say?	
7 .	When you have a job interview, don't forget to wear a suit. What else should you remember to do?	
8 .	What is your favorite holiday or festival? Tell me about it.	

9 .	If people only work four days a week, their lives will be better. What do you think?	
10.	Tell me one thing you are good at and one thing you are not good at.	
	Total	/20
	Pass?	Yes/No

Level 2 Questions

#	Questions	Points
1 .	I like people who have a good sense of humor. What kind of people do you like?	
2 .	Do you think being a teacher or a salesman is better? Give me two reasons why.	
3 .	If you want to borrow this pen, what do you say?	
4 .	Tell me about something that happened on your way to work or school last week.	
5 .	What are two things visitors should do when they go to someone's house in your country?	
6 .	Tell me about a time you had a problem with something you bought.	
7 .	What can be done to help the homeless or to help the environment?	
8 .	Would you rather study English in Australia or Canada? Would you prefer to take a painting or photography class?	
9 .	How I can meet new people and make new friends?	
10.	What kind of mobile phones do you think everyone will be using in the future?	
	Total	/20
	Pass?	Yes/No

Level 3 Questions

#	Questions	Points
1 .	What is one thing you can't stand doing?	
2 .	What is the best thing about being able to speak English?	
3 .	How can we make schools in this country better?	
4 .	What are two good ways to learn a foreign language?	
5 .	Tell me about a city you have enjoyed visiting.	
6 .	What do you do when you need some energy?	
7 .	What do you do if you can't sleep at night?	
8 .	Generally, what kind of music do your friends listen to?	
9 .	What are some good ways to start a conversation?	
10.	What kinds of things do you talk about with your spouse/family that you don't discuss with friends?	
	Total	/20
	Pass?	Yes/No
	SUGGESTED LEVEL ____	

Appendex C: Final Level Placement Chart

Part A Suggested Level	Part B Suggested Level	Placement Level
Level 1 or lower		suggest lower-level course
Level 1	1	L 1
Level 2		L 1
Level 3 or higher		L 2
Level 1 or lower		L 1
Level 1	2	L 2
Level 2		L 2
Level 3 or higher		L 3
Level 1 or lower		L 2
Level 1	3	L 3
Level 2		L 3
Level 3 or higher		suggest higher-level course