

An alternative to the traditional interview test: The observed pair interview

Armando Duarte
mando.duarte@gmail.com
University of Southern California

Abstract

This paper will report on a performance-based assessment which was performed in a public junior high school setting in west Japan. After researching some of the difficulties in implementing performance-based speaking assessments, the author devised an alternatively formatted assessment for a small group of students. The target group of students, assessment placement in curriculum, assessment design, rubric and scoring, and limitations will be discussed.

Keywords: speaking, performance-based, alternative

Veteran practitioners of English will be well aware of the difficulties that accompany any assessment design and implementation. The interplay between authenticity, practicality, reliability, validity, and washback is a balancing act which must be attended to with great care. The problems with speaking tests, in particular, have been well-documented in research (Brown & Abeywickrama, 2010; Hirai & Koizumi, 2009), with the most common difficulties found being the burden of carving out time to speak with individuals or small groups of students (if the assessment calls for such small groupings) and the relationship between speaking and either listening, reading, or writing which can detract from creating a “pure” speaking test. This speaking assessment was designed with the intent to increase a traditional interview test’s practicality and authenticity as much as possible.

Student population

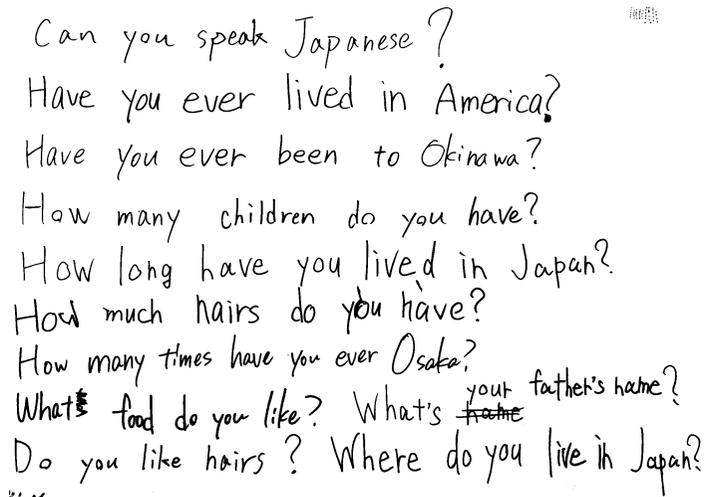
The students who underwent this test are a group of 21 Japanese L1 speakers attending a public junior high school in the city of Kakogawa, Hyogo. This assessment was performed in the spring of 2016 when the students had just begun their third year of study at the secondary school level. Combined with two years of instruction during the last two years of secondary school, the students had been studying English as a foreign language for four academic years. As English is mandatory in junior high school, the students’ level of motivation varied widely; ability-wise, most students would be in the “expanding” level of English proficiency according to the California ELD standards (California Department of Education, 2014), meaning that they can use English to engage in basic communication with others on topics that are familiar to them but they have yet to use English for academic purposes.

Assessment in the curriculum

This assessment was performed at the end of a three-unit task-based lesson centered around using the present perfect tense to elicit information from a partner. In the first lesson, the students watched a short self-introduction and formulated questions in groups to ask this person. These questions were collected on group posters and combined to form a “bank” of questions which were used in later activities. The second lesson saw the students creating a list of six questions, three using the present perfect and three using either the simple past or simple present, to interview a classmate. These questions were written, rewritten, and the interview was carried out all in the second lesson. Students presented their findings and engaged in active listening by reporting one new fact they had learned about their classmates during the presentations.

Assessment design and procedure

In keeping with the principle of content validity, where students should be tested on what they have already performed in class, the questions for this assessment were taken from the “bank” of questions that the students themselves generated during the first lesson. Every individual question was typed out and used as the basis for the test questions. The most common patterns and lexical items were lifted out of this list of questions. Any grammatical or spelling errors present in the students’ original questions were corrected as they were entered into the “bank” – to ensure consistency during the actual test - but a review of the questions students created (see Figure 1) indicate that errors were minimal.



Can you speak Japanese?
 Have you ever lived in America?
 Have you ever been to Okinawa?
 How many children do you have?
 How long have you lived in Japan?
 How much hairs do you have?
 How many times have you ever Osaka?
 What ~~is~~ food do you like? What's ~~the~~ ^{your} father's name?
 Do you like hairs? Where do you live in Japan?

Figure 1. Example of student-generated questions

The observed interview format itself came as a result of previous research indicating the impracticality of interview tests. The unified oral testing format (Ramirez, 2014), in which students create role plays with randomly-assigned partners and then engage in a question-and-answer session, was used as a starting point and modified to fit the parameters which were present in this particular classroom context. The academic calendar demanded that this assessment be implemented and finished before the spring holidays, so the role play was discarded and the resulting assessment was what was eventually carried out – an assessment where one student interviews another and is then interviewed in turn. Students did not know who their partner would be until they were called to the testing area (the hallway outside the classroom) but they were told the questions used for the assessment would come from the “bank” of questions they had previously created.

Students were told to enter the testing area two at a time and sit down. Student A was instructed to pick up his or her sheet of paper and ask student B the questions which appeared on said sheet. Student B was told to respond. After student A’s questioning had run its course, student B was told to pick up his or her paper and ask the questions that appeared. Student A was told to respond. Students were dismissed at the end of this Q&A session. Two new students entered and the cycle repeated itself.

Rubric and scoring

Given that each student is playing the role of interviewer and interviewee, two rubrics were created for this assessment, as shown in Figure 2 and Figure 3. As the interviewer, students were assessed on their ability to ask – that is, read – questions without errors, as shown in Figure 4. Students playing the part of the interviewee were assessed on their ability to answer questions in two full, grammatically correct

sentences, which they had been instructed to do before the start of the assessment. This instruction took the form of a role play between the native and Japanese instructors. In rare cases where interviewee students answered in more than two sentences, extra contributions which included errors did not count against them.

Numeric score	Evaluation
3	I can understand and answer questions with no mistakes using 2 sentences.
2	I can understand and answer questions with mistakes in 1 or 2 questions using 2 sentences <u>or</u> I said “no” to question 3 (Okinawa).
1	I can understand and answer questions with mistakes in 3 or 4 questions using 2 sentences <u>or</u> I skipped a question <u>or</u> I didn’t use 2 sentences
0	I could not answer any questions or I asked the teacher for help.

Figure 2. Interviewee rubric

Numeric score	Evaluation
3	I can read and ask questions with no pronunciation mistakes.
2	I can read and ask questions with mistakes on 1 or 2 questions.
1	I can read and ask questions with mistakes on 3 or 4 questions.
0	I could not ask any questions or I needed help.

Figure 3. Interviewer rubric

It would not be enough to simply say “Yes” to the question “Have you ever been to Okinawa” – students would have to answer “Yes, I have. I have been to Okinawa” to receive credit for accurately answering the question. The Okinawa question itself was important because all students taking this assessment had recently returned from a school trip to the island, so any negative answer to that question was judged as students giving any answer to satisfy the question. Otherwise, interviewee students were assessed on answering in complete sentences and using parallel verb forms from the question in the answer – “Do you know how to play the violin” requires that students also use “know how to” in their response, for example. Prior to the test, students were instructed to give their answers in this manner and responses that did not fit this criterion were penalized even if a native speaker would understand the response. This choice was made in keeping with the test’s overall goal of accuracy.

The results from this assessment, shown in Figure 5, indicate that almost every student is able to read and ask questions accurately. However, students as interviewees are not able to answer questions more accurately when being asked by their peers compared to a traditional interview test with a native English teacher. In other words, students did not perform better just because they were being interviewed by their classmates but rather, they still made the same mistakes they might otherwise make. This, to me, is actually a redeeming feature of the test in that it indicates that the native teacher can be swapped out for

a student and the resulting answers will be identical or nearly identical.

A

Hello. How are you?			
Do you have any pets?	Can you play the piano?	Have you ever been to Okinawa?	Have you ever eaten Chinese food?
Thank you.			

B

Hello. How are you?			
Do you like soccer?	Can you speak English?	Have you ever been to Okinawa?	Have you ever played tennis?
Thank you.			

Figure 4. Assessment questions

	Interviewer Score	Interviewee Score	Comments
Tanimoto Yuma	3	1	Good job! Try to answer in 2 sentences.
Fujiwara Nanako	3	2	Very nice! Remember: "Yes, I have. I have eaten ---" not "Yes, I have. I have never eaten ---"
Maeda Kana	3	2	Try to ask and answer questions faster.
Nagaoka Takuma	2	2	Try to speak faster: asking and answering.
Nagatani Yuki	2	2	Remember "pets" not "petos" Good speed "Chinese food" not "China's food" Good self-correction!
Masaki Airi	3	2	Asking for clarification – fantastic! Answer in 2 sentences Good asking speed
Taniyama Kenta	3	1	Try to ask faster Try to answer in 2 sentences, and don't skip a question! Try your best.
Fujimoto Yuuki	3	2	"I haven't pets" is "I don't have a pet" Try to speed up a little
Tamura Aika	3	2	Good asking speed "I have ever been to" is "I have never been to"
Okamura Yuma	2	2	Fantastic! Speed, correction, pronunciation
Nagaoka Mayuki	2	2	Answer in 2 sentences: "Yes, I have. I have been to Okinawa."
Tanimoto Gen	3	2	Use a bigger voice "a little" – good answer!

Figure 5. Student scores and feedback. (Student names have been changed in the interest of privacy.)

Limitations and conclusion

Several limitations were made apparent from the moment the first pair sat down to perform their observed interview. The first student A picked up his question sheet and began to ask all three questions back to back, without giving his partner a chance to answer. This resulted in the instructor having to intervene and ask the student to ask one at a time, which itself was something of a problem. Various students, as interviewees, were unsure of what to do when asked a question by their peers. This resulted in several blank stares at the instructor, which, at least once, prompted the instructor to provide a binary option – "Yes? No?" which then allowed the interviewee to proceed.

The shortcomings of traditional interview tests do not entirely disappear with this assessment. Although students are being assessed two at a time and the rubric gives clear guidelines for grading, the instructor must follow the interaction between the students extremely closely not only to judge for accuracy but also to look for areas of praise and improvement – information which is then passed on to the student. When this assessment was actually carried out, this aspect was the most challenging as there was very little time between the exit of one pair and the entrance of the next. Future attempts at using this assessment could

be improved by instructing students to wait a minute or so before entering the testing area after the preceding pair has exited.

Overall, this assessment provides several benefits compared to a typical interview test. It improves on the practicality of a traditional interview test by allowing the instructor to assess two students at once, while the inclusion of a peer can help lower students' affective filters, decreasing overall nervousness. The testing format is authentic in that asking a conversation partner about themselves is an act that occurs in the real world, not just in the classroom. It is easy to imagine these same junior high school students asking similar questions if they find themselves speaking to an exchange student, for example. However, the scripted format of the conversation reduces the authenticity and validity of the assessment. If this test is seen as a test of whether or not students can have a natural conversation, clearly the script in front of the test-takers (and the reading of questions) reduces validity in that sense, no matter how authentic the questions. Given that students are participating in an activity that closely mirrors a classroom activity using language that they themselves formulated, some measure of validity is regained. The rubric, meanwhile, provides clear, reliable guidelines for grading and comments provided to students give them some direction for improvement – hints that will help them answer correctly in the future or simply advice on how to sound more natural when speaking.

However, given the close relationship between the four major skills of language, further improvements can be made. Cued storytelling, where students are presented with visual input which drives speaking, could be integrated into this assessment. Students could be informed that they will be judged on their ability to ask questions in the present perfect, for example, then given cards with pictures depicting people engaging in various activities. This format more closely mirrors a “pure” speaking test in that students do not have to rely on their reading abilities, as they do in the current assessment, to ask questions. This modification could improve on the current assessment's validity – the most important aspect of assessment.

References

- Brown, H. D., & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices* (2nd ed.). White Plains, NY: Pearson Education.
- California Department of Education. (2014). *California English language development standards: Kindergarten through Grade 12*. Retrieved from <http://www.cde.ca.gov/sp/el/er/documents/eldstndspublication14.pdf>
- Hirai, A., & Koizumi, R. (2009). Development of a practical speaking test with a positive impact on learning using a story retelling technique. *Language Assessment Quarterly*, 6(2), 151-167. doi: 10.1080/15434300902801925
- Ramirez, C. (2014). Balancing validity, practicality and reliability on an university oral English test. *Kinki University Center for Liberal Arts and Foreign Language Educational Journal*, 5(1), 73-95. Retrieved from kurepo.clib.kindai.ac.jp/modules/xoonips/download.php?file_id=11458