Creating a Listening Test to Track Young Learners' Language Development

Patrick McLain, Brent Miller, Renée Saulter, & Rachele Stucker, Michigan Language Assessment

Introduction

We share the results of two pilot studies to highlight the performance of multi-level listening task types for a test for young learners, and show how our predictions of the ability levels assessed by each task type were reaffirmed in our results. We also discuss design considerations made to enhance accessibility and test taker engagement with the test.

MET Go!

- full-color, four-skills test designed for 11- to 15-year-olds
- designed to track young learners' development in English
- multi-level: targets A1 to B1 on the CEFR

MET Go! Listening Test

- consists of five item types (six were piloted)
- an application of Buck's (2001) "default listening construct"

Designed for Young Learners

- stimuli are double played
- content focused on educational, public, and personal domains
- test items are similar to classroom tasks

Skills Commonly Assessed in Language Tests

more appropriate for young learners

- understand main idea/gist
- listening for specific details
- understand vocabulary
- identify speaker's attitude/ purpose

less appropriate for young learners

- make an inference/conclusion
- understand rhetorical function/pragmatic implications
- synthesize ideas from different parts of the text

Sample Item

Figure 1 features a sample of the announcement item type from MET Go! Accessibility to test takers and common classroom tasks were considered in the design of listening test item types.

Figure 1: Announcement

[audio only]

- N: Listen to the talks. You will hear each talk twice. You can take notes while you listen. Then answer each question.
- N: Listen to a teacher talking to her class.
- W: It's almost time to leave for our field trip.
 As you know, today we're going to a furniture factory. We will watch furniture, like tables, chairs, and desks, be made.

When we get back to school, I want you to write a report. In it you should describe what you saw at the factory. You can write about the machines and what the workers were doing.

OK, please make a line by the door. The bus we'll take to the factory will be arriving in just a couple of minutes.

- N: What is the woman mostly talking about?
- N: What does the woman say about tables and chairs?

[printed in test booklet]

Listen to the talks. You will hear each talk twice. You can take notes while you listen. Then answer each question.



- 1. What is the woman mostly talking about?
- A. new classroom furniture
- B. something she saw on a trip
- C. today's class activity
- 2. What does the woman say about tables and chairs?
- A. The class will see them at the factory.
- B. The students should sit on them.
- C. The class has new ones.

the MET Go! Listening test construct.

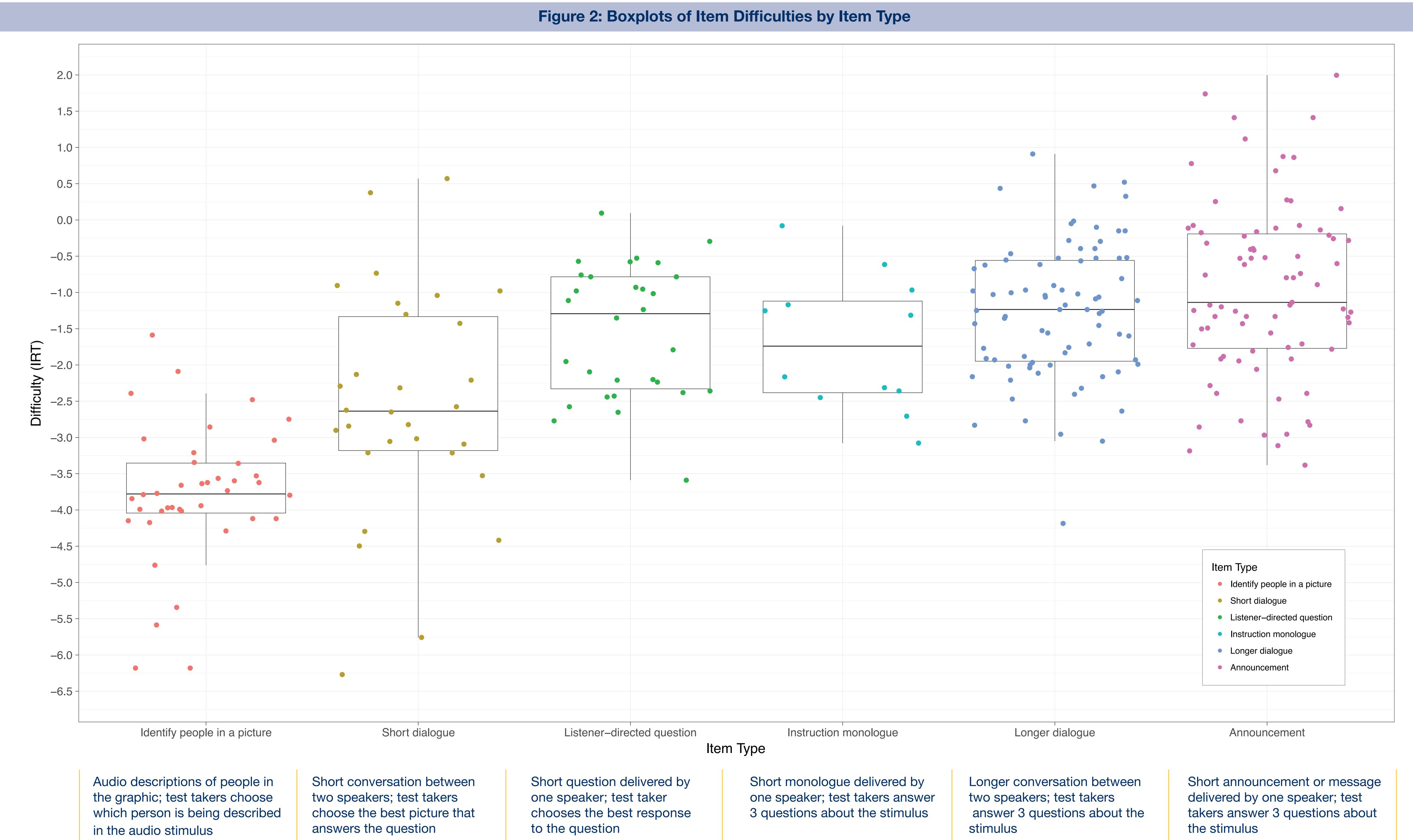
Pilot Study Design

• In the spring of 2017 we began development of

- The pilot tests were administered in late 2017 and mid-2018.
- 270 listening items were piloted across 10 forms in 10 countries and approximately 2,700 test takers.

Pilot Test Results

Results of the two pilot tests were very similar. Figure 2 is an aggregate of all item performance data from both pilot tests.



Changes Made to the Listening Test

Overall, the item types that were developed performed as expected, with the identify people in a picture and short dialogues being less difficult than the listening question, longer dialogue, monologue, and announcement item types.

Target: A1

We made the following changes to the item types post-pilot:

Target: A1-B1

- removed instruction monologue item type: did not perform at anticipated difficulty level
- increased stimulus length of longer dialogue and announcement item types: to allow for one more question per stimulus

Next Steps

Target A2-B1

The MET Go! test development team is currently finalizing the process for delivery of test taker performance feedback via a score report. The nature of our multi-level listening test (i.e., multiple item types targeting a variety of difficulty levels and skills) allows us to send comprehensive feedback to young learners about their performance.

References

Buck, G. (2001). Assessing listening. Cambridge: Cambridge

Council of Europe. (2011). Common European Framework of Reference for Languages: Learning, Teaching, Assessment.

Field, J. (2015). The effects of single and double play upon listening test outcomes and cognitive processing. British Council.

Target: A2-B1

Gu, L. (2015). Language ability of young English language learners: Definition, configuration, and implications. Language Testing, 32(1), 21-38.

Zieky, M. J., Perie, M., & Livingston, S. A. (2008). *Cutscores:* A manual for setting standards of performance on educational and occupational tests. Princeton, NJ: Educational Testing Service.

Contact Information

Brent A. Miller, Assessment Developer Michigan Language Assessment Email: miller.b@michiganassessment.org

Target: A1-B1



Target: A2-B1