



# MICHIGAN LANGUAGE ASSESSMENT

## MET Go!

### **Development of the MET Go! Listening and Reading Tests**

Technical Report

## **Contact Information**

---

All correspondence and mailings should be addressed to:

### **Michigan Language Assessment**

Argus 1 building  
535 West William St., Suite 310  
Ann Arbor, Michigan  
48103-4978 USA

T +1 866.696.3522

T+1 734.615.9629

F +1 734.763.0369

[info@michiganassessment.org](mailto:info@michiganassessment.org)

[MichiganAssessment.org](http://MichiganAssessment.org)



© 2019 Cambridge Michigan Language Assessment.

**MICHIGAN  
LANGUAGE  
ASSESSMENT**

 Cambridge Assessment  
English

 UNIVERSITY OF MICHIGAN

## Table of Contents

1. Introduction .....	1
2. Test Construct.....	1
2.1 Targeted Levels in the CEFR.....	1
2.2 Listening Construct Definition.....	2
2.3 Reading Construct Definition .....	3
3. Test Development .....	4
3.1 Test Design .....	4
3.2 Pilot Testing.....	4
3.3 Personalized Feedback .....	5
4. Interpreting Listening and Reading Test Scores.....	7
5. References .....	8

## List of Tables

Table 1: Overall Listening Comprehension (Council of Europe, 2018) .....	2
Table 2: Overall Reading Comprehension (Council of Europe, 2018) .....	2
Table 3: MET Go! Listening Test Parts, CEFR Levels Targeted, and Item Descriptions .....	4
Table 4: MET Go! Reading Test Parts, CEFR Levels Targeted, and Item Descriptions.....	4
Table 5: MET Go! Listening Personalized Feedback Comparison .....	7

## List of Figures

Figure 1: MET Go! Listening Pilot Item Difficulty Boxplots.....	5
Figure 2: MET Go! Reading Pilot Item Difficulty Boxplots .....	6
Figure 3: Sample MET Go! 4 Skill Score Report Personalized Feedback Boxplots.....	6
Figure 4: Diagram of MET Go! Listening and Reading Descriptor Assignment Procedures .....	7

## 1. Introduction

The MET Go! is a multi-level test of English language ability designed for beginner to intermediate level learners of middle and secondary school age. Developed and produced by Michigan Language Assessment, the test covers the four language skills (listening, reading, speaking, and writing), assessing learners' ability in each area and assisting them as they progress in their learning.

The MET Go! Listening and Reading Tests are designed to assess test takers' English listening and reading proficiency by evaluating their ability to understand spoken and written English on a variety of familiar school and everyday topics across several different item types. For the listening test, the items include asking test takers to identify people in a picture based on an audio description, answering questions asked by a single speaker, answering questions based on both short and long conversations between two speakers, and answering questions based on a short announcement or message delivered by a single speaker. Audio for the recorded information and questions are played twice. For the reading test, the items include answering grammar and vocabulary related questions, as well as answering questions in relation to informational and narrative texts.

The MET Go! Listening and Reading Tests are intended to be useful in a variety of educational settings. The results can be used to monitor the progress of English as a Second Language (ESL) or English as a Foreign Language (EFL) learners, as well as for placement or diagnostic purposes to inform instructors of the strengths and weakness of the learners and areas where instruction is needed. Language programs can also use the test to certify whether or not learners have achieved the goals of a language course.

This report describes the development of the MET Go! Listening and Reading Tests. It provides information on the development of the test construct and task types, as well as information on score interpretation.

## 2. Test Construct

### 2.1 Targeted Levels in the CEFR

The Common European Framework of Reference (CEFR) provides a common basis for evaluating the ability level of language learners. The framework offers illustrative scales and can-do statements that describe "what language learners have to learn to do in order to use a language for communication and what knowledge and skills they have to develop so as to be able to act effectively" (Council of Europe 2001, p. 1).

The MET Go! Listening and Reading Tests were specifically developed to assess test takers' listening and reading abilities at the A1-B1 levels of the CEFR. Both the original CEFR volume (Council of Europe, 2001) and the more recent companion volume (Council of Europe, 2018) were used by the MET Go! Listening and Reading Test development teams throughout the development process as references to inform the design of the test construct, the item types, and the overall test design.

The can-do statements from numerous CEFR illustrative scales were heavily referenced during development. For the listening section these scales included the overall listening comprehension, understanding conversation between other speakers, and listening to announcements and instructions scales (Council of Europe, 2001, 2018). For the reading section these scales included the overall reading comprehension, reading correspondence, and reading for information and argument scales (Council of Europe, 2001, 2018). Tables 1 and 2 summarize the progression in overall listening comprehension and overall reading comprehension, respectively from levels A1 to B1 for learners aged 11 – 15 (Council of Europe, 2001, 2018). As learners progress through each CEFR level they are expected to have mastered abilities described under lower levels of competence. The tables show that A1 level test-takers are able to understand very slow and carefully articulated speech on familiar topics and very short, simple texts a single phrase at a time. More proficient test-takers are able to understand the main points of clear, standard

**Table 1: Overall Listening Comprehension (Council of Europe, 2018)**

CEFR Level	Descriptor
B1	Can understand straightforward factual information about common everyday or job related topics, identifying both general messages and specific details, provided speech is clearly articulated in a generally familiar accent.
	Can understand the main points of clear standard speech on familiar matters regularly encountered in work, school, leisure etc., including short narratives.
A2	Can understand enough to be able to meet needs of a concrete type provided speech is clearly and slowly articulated.
	Can understand phrases and expressions related to areas of most immediate priority (e.g. very basic personal and family information, shopping, local geography, employment), provided speech is clearly and slowly articulated.
A1	Can follow speech that is very slow and carefully articulated, with long pauses for him/her to assimilate meaning.
	Can recognise concrete information (e.g. places and times) on familiar topics encountered in everyday life, provided it is delivered in slow and clear speech.

**Table 2: Overall Reading Comprehension (Council of Europe, 2018)**

CEFR Level	Descriptor
B1	Can read straightforward factual texts on subjects related to his/her field and interests with a satisfactory level of comprehension.
A2	Can understand short, simple texts on familiar matters of a concrete type which consist of high frequency everyday or job related language.
	Can understand short, simple texts containing the highest frequency vocabulary, including a proportion of shared international vocabulary items.
A1	Can understand very short, simple texts a single phrase at a time, picking up familiar names, words and basic phrases and rereading as required.

speech on familiar matters and read straightforward factual texts on subjects related to his/her field and interests.

## 2.2 Listening Construct Definition

The MET Go! Listening Test construct is based on Buck's "default listening construct" (2001). This construct is defined as the ability to process extended samples of realistic spoken language, automatically and in real time, to understand the linguistic information that is unequivocally included in the text, and to make whatever inferences are unambiguously

implicated by the content of the passage (Buck, 2001, p. 114).

The MET Go! Listening Test construct is modified slightly from Buck because the proficiency of the target test taking population does not extend to the levels that are covered by his default construct. Considering the age and expected proficiency of target test-takers, the construct was narrowed to the ability to comprehend straightforward and factual input that is delivered in standard, clearly articulated American English in a variety of familiar settings that beginner to intermediate level

learners of middle and secondary school age might encounter in the course of routine daily/school life (e.g., classrooms) and on a variety of general interest (i.e., non-specialized) topics that are frequently encountered in those settings. "Comprehend" in this case refers to the following abilities:

- understand the communicative intent of short exchanges and talks
- extract main ideas and gist from short dialogues and monologues
- generally follow the most relevant and salient details contained within the input
- recognize words in short, descriptive statements
- understand level-appropriate vocabulary

Some abilities commonly assessed in language tests were considered to be less appropriate for the target test population, and were therefore excluded from the listening test construct. These abilities included:

- making inferences/conclusions
- understanding rhetorical function/pragmatic implications
- synthesizing ideas from different parts of text.

Additionally, in order to ensure that the listening test was appropriate for learners age 11 to 15: audio input is played twice (cf. Field, 2015; Ruhm et al., 2016; and Goodwin, 2017), listening items feature engaging full-color graphics, and test items are similar to tasks completed in the classroom. It is hoped that these features create a positive assessment experience for these younger learners.

### 2.3 Reading Construct Definition

The development of the MET Go! Reading Test construct was heavily influenced by Hasselgreen and Cauldwell (2016), which utilized Khalifa & Weir's (2009) socio-cognitive model of reading. This model primarily consists of four parts: "text/input, features of the task (types of reading), knowledge and processing" (Hasselgreen & Cauldwell, 2016, p.56), and it views the reading construct as "residing in the interactions between the underlying cognitive ability, the context of use and the process of scoring" (Khalifa & Weir, 2009, p.3). The MET

Go! Reading Test construct also considers the can-do descriptors of the relevant CEFR scales at the A2-B2 levels, as well as the first language literacy and overall cognitive linguistic development of language learners aged 11-15. While A1 readers are extremely restricted in terms of what tasks are "doable" (Hasselgreen & Cauldwell, 2016, p. 69) – typically only handling single-word texts – those at age 11 and above and level A2 and above can attempt tasks drawing on nearly all aspects of a model of reading. For example, in Khalifa and Weir's (2009) model of reading, they should be able to read both carefully (detail-oriented) or quickly (skimming) and understand a text globally or locally. Nevertheless, texts and tasks must not be overly long, abstract, or require sophisticated logical analysis; processing texts of that nature can typically only be expected of young adults (17+) at the C1 level or higher.

The construct of the reading test therefore covers several skills and abilities that readers can perform on texts ranging from one sentence to a few paragraphs, dealing with matters that are engaging and understandable to students aged 11-15. These skills include:

- recognizing/identifying words and phrases within a text
- selecting words with appropriate lexical meaning
- selecting words or phrases that are syntactically well-formed
- understanding propositional meaning within a text
- understanding the gist of a text
- understanding the author's attitude or opinion
- interpreting written dialogue

Additionally, in order to ensure that the reading content is appropriate for the age of the test takers, all content is in the educational, public, or personal domains, and deals with topics that frequently occur in these settings (e.g. social events, school assignments, or simple transactions).

### 3. Test Development

#### 3.1 Test Design

The MET Go! Listening and Reading Tests are paper-and-pencil test of English listening and reading ability that consists of 60 selected response questions (30 listening; 30 reading). Each section consists of multiple parts designed to be accessible to both lower- and higher-level test-takers (CEFR levels A1-B1). Tables 3 and 4 describe the general format of the listening and reading section, respectively, the purpose of each test part, the CEFR levels targeted, and the corresponding linguistic functions.

#### 3.2 Pilot Testing

Pilot testing was conducted throughout the development of the MET Go! Listening and

Reading Tests to ensure that they functioned as intended. In total, 12 pilot test forms were administered to 3,413 test takers from 10 countries (Argentina, Bolivia, Brazil, Colombia, Greece, Mexico, Paraguay, Peru, South Korea, and Uruguay) across three rounds of pilot testing. Test taker response data were used to obtain information on the overall performance of the test, the item types, and the individual items to help inform decisions made by the development team. Additionally, surveys were conducted throughout the different rounds of pilot testing to obtain feedback from test takers and administrators on a variety of different aspects of the test.

As a result of information gained from pilot testing, a couple changes were made to some of the listening and reading item types. For the

**Table 3: MET Go! Listening Test Parts, CEFR Levels Targeted, and Item Descriptions**

Test Part	Item Type	Levels Targeted	Description
Part 1	Identify people in a picture	A1	Audio descriptions of people in the graphic; test takers choose which person is being described in the audio stimulus
Part 2	Short dialogue	A1 – B1	Short conversation between two speakers; test takers choose the best picture that answers the question
Part 3	Listener-directed question	A1 – B1	Short question delivered by one speaker; test taker chooses the best response to the question
Part 4	Longer dialogue	A2 – B1	Longer conversation between two speakers; test takers answer 3 questions about the stimulus
Part 5	Announcement	A2 – B1	Short announcement or message delivered by one speaker; test takers answer 3 questions about the stimulus

**Table 4: MET Go! Reading Test Parts, CEFR Levels Targeted, and Item Descriptions**

Test Part	Item Type	Levels Targeted	Description
Part 1	Grammar and Vocabulary	A1 – B1	An incomplete sentence is followed by a choice of words or phrases to complete it. Only one choice is grammatically correct or has the correct meaning in that context.
Part 2	Reading Comprehension	A1 – B1	Short texts with multiple questions per passage. The first two texts are brochures, emails, advertisements, announcements, or other similar materials. The second two texts are short narrative stories.



listening test, item performance data from the first round of piloting resulted in the removal of one item type, instruction monologue. It had consisted of 3 questions on a short monologue delivered by a single speaker. The item type did not perform at the anticipated difficulty level and overlapped with several of the other listening item types, so it was ultimately not included in the final test specifications. The stimulus length of the longer dialogue and announcement item types were also increased following the first round of pilot testing to allow for one more question per stimulus. For the reading test, item performance data prompted some changes to the item writer guidelines for the reading comprehension items to widen the range of target difficulty levels so that the items more accurately reflect the intended design. Additionally, the art specifications for the informational reading comprehension items were slightly modified so that they would be more representative of real-life media.

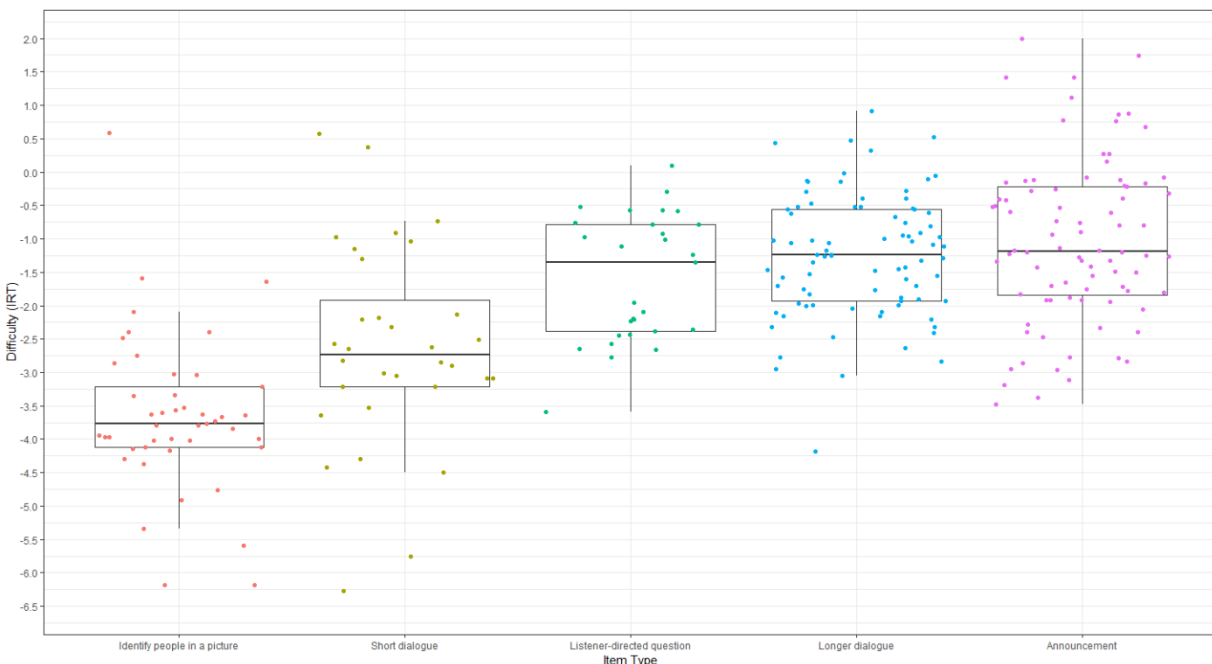
Overall, the three rounds of pilot testing showed that the different MET Go! Listening and Reading Test items generally performed as expected, with each of the listening and reading item types covering a wide range of difficulties. Figures 1 and 2 present boxplots of the item

difficulties by item type for the listening and reading tests, respectively. Figure 1 shows that for the listening test the identifying people in a picture and short dialogue items were generally less difficult than the listener-directed question, longer dialogue, and announcement items for the listening test, and Figure 2 shows that for the reading test the grammar and vocabulary and reading comprehension item types covered similar ranges of item difficulties. These different item difficulty distributions are comparable to the target CEFR levels of each item type summarized in Tables 3 and 4, which suggests that the MET Go! Listening and Reading Test items functioned as expected.

### 3.3 Personalized Feedback

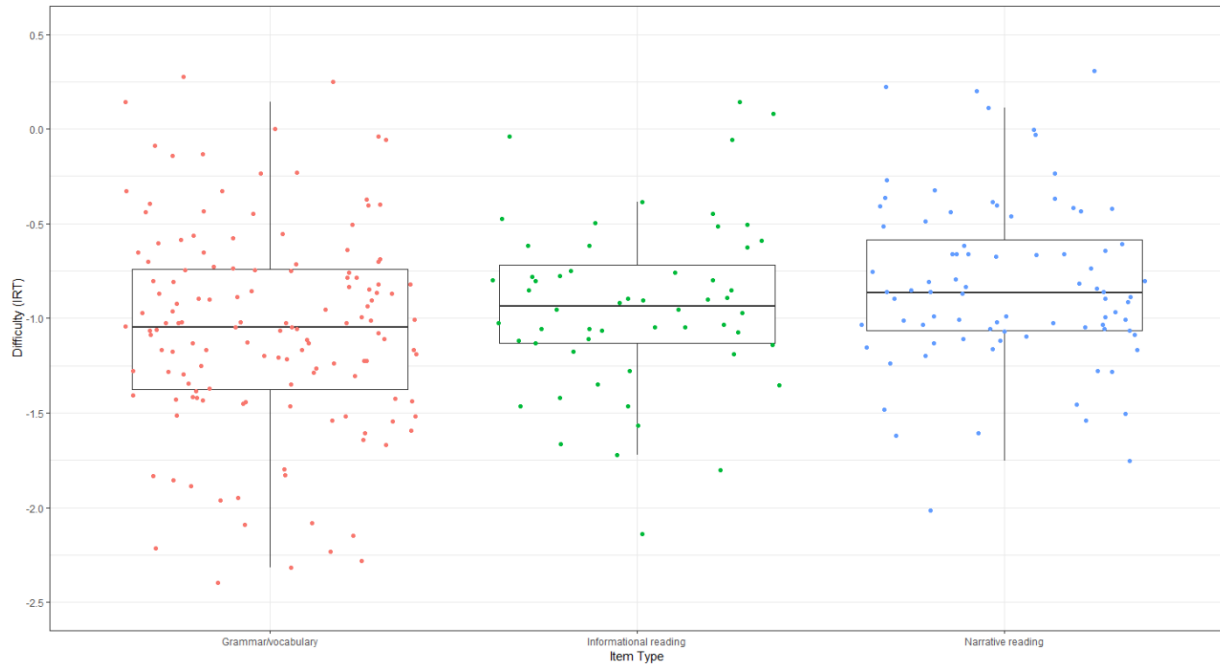
Providing test takers with scores is important because they provide a reliable evaluation of test takers' overall ability level; however, scores alone are insufficient if the aim of the test is to have an impact on learning (Alderson, 2005; Kunnan & Jang, 2009; Lee, 2015). One of the key features of the MET Go! is that it provides test takers with personalized feedback based on their individual performance to help impact learning by highlighting what they did well and what they can improve upon,

**Figure 1: MET Go! Listening Pilot Item Difficulty Boxplots**





**Figure 2: MET Go! Reading Pilot Item Difficulty Boxplots**



as well as suggesting learning activities that they can do to improve.

Figure 3 provides a sample of the kind of personalized feedback that test takers who take the 4-skill MET Go! can expect to receive on their score report. The MET Go! provides test takers with both personalized feedback descriptors and recommended learning activities for each test section. Personalized feedback descriptors are aimed at helping to provide test takers with an understanding of their strengths and weaknesses to give them more detailed information on their performance and what they need to improve, while recommended learning activities are meant to provide test takers with interesting, authentic activities to help test takers pursue learning on their own. The language used in the personalized feedback was written at the level of English ability that they represent in order to ensure that the feedback was accessible to each test taker.

For the MET Go! Listening and Reading Tests, every item is tagged based on the different sub-skills that they test, which provides the basis for the personalized feedback. For listening, the sub-skills evaluated include vocabulary knowledge, grammar knowledge, extended speech comprehension, local detail

recognition, main idea recognition, and identification of the speaker's purpose/attitude. For reading, the sub-skills evaluated include vocabulary knowledge, grammar knowledge, local detail recognition, main idea recognition, and the ability to synthesize information from more than one sentence. The personalized

**Figure 3: Sample MET Go! 4 Skill Score Report Personalized Feedback Boxplots**

**MET Go!**

**Continue Your English Language Learning**

We looked carefully at your answers on the test so we can tell you more about your English. Below is information about what you are good at, what to improve, and things you can do next. Keep learning!

**About Your English**

**LISTENING** You can understand conversations and talks about everyday things well. You know enough English words and grammar to follow simple talks and conversations. However, work on listening for how speakers feel about the topic and the reason they're talking, too.

**READING** You can understand some of the details in short, simple readings, and sometimes find the main idea or purpose, too. Sometimes you know how sentences connect to each other as well. You know some basic English words, and they sometimes help you understand the things you read. Your grammar knowledge also helps you to understand simple sentences and some longer ones.

**WRITING** You can tell a story when you write. Next, try writing about your personal experiences and opinions. Also, you could add more details to your writing to make it more interesting. To help readers follow your writing, try using more connecting words. Grammar and vocabulary mistakes sometimes make your writing hard to understand.

**SPEAKING** You can describe familiar places and things in detail. However, try to connect your ideas when you talk about what people do. When you talk about a personal experience, you include a lot of interesting information. You're also able to explain why you believe something, but make sure you say clearly what your opinion is. You sometimes try to use more advanced grammar and longer words, too. Keep it up! Finally, your message is usually clear when you talk.

**Things To Do**

Watch reviews of movies or products you like. What is the reviewer's opinion? How do you know?

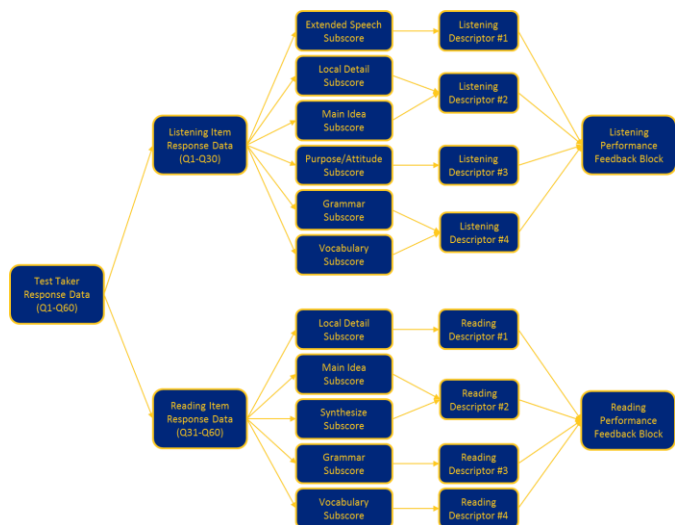
Read an English website about your favorite actor, athlete, or musician. Use a dictionary to help you understand new words. Tell a classmate or family member about what you learned.

Read some pages from an English story. Look for words like "and," "then," "but," and "because." Now, write in English about the things you did today. Can you use one of those words in every sentence?

Try going to a store with a friend or family member. Find a product that you like or don't like. Use English to explain why.

feedback descriptors and recommended learning activities are assigned based on the test takers' performance on these different subscores. Figure 4 provides a diagram to help visualize the assignment procedures for this personalized feedback.

**Figure 4: Diagram of MET Go! Listening and Reading Descriptor Assignment Procedures**



Because the personalized feedback depends on a number of different factors related to the test takers performance, test takers with similar listening and reading scores may in fact receive different personalized feedback based on the sub-skills tested by the items they answered correctly or incorrectly. Table 5 presents an example that compares the personalized feedback received by two test takers (A and B) with identical scores on the MET Go! Listening

Test. It shows that while both test takers performed well on items testing extended speech comprehension and received the same initial descriptor, the rest of the descriptors are different. These difference reflect the fact that the two test takers performed differently on items testing local detail and main idea recognition, identification of the speaker's purpose/attitude, and grammar ability.

## 4. Interpreting Listening and Reading Test Scores

MET Go! Listening and Reading Test scores are intended to reflect test takers' ability to listen and read in English. Test takers who complete the listening and reading tests will receive a score report that includes a scaled score (0-52) and CEFR level (Below A1-B1) for each section based on their overall listening and reading performance, as well as personalized feedback in the form of a performance descriptor statement and a recommended learning activity based on their performance on the different parts of the listening and reading tests. For test takers, these results can help them to recognize their strengths and weaknesses and decide on strategies for improving their English. For ESL/EFL instructors, these results can help them place students into appropriate classes, monitor the progress of students in a class, and provide diagnostic information to identify areas where instruction is needed.

**Table 5: MET Go! Listening Personalized Feedback Comparison**

Test Taker	Scaled Score	CEFR Level	Feedback Descriptor
A	40	B1	You can understand conversations and talks about everyday things well. When listening, you can understand details, but try to listen for the main points, too. However, you can recognize people's attitude and purpose in conversations and talks.
B	40	B1	You can understand conversations and talks about everyday things well. When listening, you get the main points, but try to listen for the details, too. Also, think about the reasons why speakers are talking and how they feel about the topic. Learning more English grammar can help you understand more of what you hear.

## 5. References

- Alderson, J. C. (2005). *Diagnosing foreign language proficiency: the interface between learning and assessment*. London: Continuum.
- Buck, G. (2001). *Assessing listening*. New York: Cambridge University Press.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: learning, teaching, assessment*. Cambridge, UK: Cambridge University Press.
- Council of Europe. (2018). *Common European Framework of Reference for Languages: learning, teaching, assessment. Companion Volume with New Descriptors*. Retrieved from <https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989>
- Field, J. (2015). *The effects of single and double play upon listening test outcomes and cognitive processing*. British Council English Language Assessment research group. ARAGs Research reports online. [https://www.britishcouncil.org/sites/default/files/field\\_layout.pdf](https://www.britishcouncil.org/sites/default/files/field_layout.pdf).
- Goodwin, S. (2017). *Locus of control in L2 English listening assessment* (Doctoral dissertation, Georgia State University).
- Hasselgreen, A. & Cauldwell, G. (2016). *Assessing the Language of Young Learners* (British Council Monographs on Modern Language Testing). Sheffield: Equinox Publishing.
- Khalifa, H., & Weir, C. J. (2009). *Examining reading: Research and practice in assessing second language reading: Studies in language testing*, 29. Cambridge, UK: Cambridge University Press.
- Kunnan, A. J., & Jang, E. E. (2009). *Diagnostic feedback in language assessment*. In M. H. Long & C. J. Doughty (Eds.), *The handbook of language teaching* (610-625). Hoboken, NJ: Blackwell.
- Lee, Y. (2015). *Diagnosing diagnostic language assessment*. *Language Testing*, 32(3), 299-316.
- Ruhm, R., Leitner-Jones, C., Kulmhofer, A., Kiefer, T., Mlakar, H., & Itzlinger-Bruneforth, U. (2016). *Playing the recording once or twice: Effects on listening test performances*. *International Journal of Listening*, 30(1-2), 67-83.