

---

# Minimal English Test Versus the TOEIC

Masaya Kanzaki

*Kanda University of International Studies*

*kanzaki-m@kanda.kuis.ac.jp*

---

This study compares the results of the Minimal English Test (MET), a gap-filling dictation test designed to evaluate the language proficiency of English learners in five minutes, with the results of three different types of the Test of English for International Communication (TOEIC) in order to determine the degree to which the scores correlate. Participants in this study were 90 university students. They completed the MET and the TOEIC listening, reading and speaking tests, and their scores were then examined for correlations. The speaking score was correlated more strongly with the MET score than with the listening score, the reading score, and the combined listening and reading score.

本研究では、英語学習者の言語能力を5分で測定するように作られた穴埋め式ディクテーションテストであるMinimal English Test (MET、最少英語テスト)と3種類のTOEICテストとの結果を比較し、各テスト間の相関関係を検証する。本研究の参加者は90名の大学生である。彼らはMETとTOEICのリスニングテスト、リーディングテスト、スピーキングテストを受験し、彼らの得点に対して相関関係の検証が行われた。スピーキングの得点は、リスニングの得点、リーディングの得点、およびリスニングとリーディングの合計点との間よりもMETの得点との間に強い相関があった。

---

Some Japanese universities administer the Test of English for International Communication (TOEIC) or the paper-based version of the Test of English as a Foreign Language (TOEFL) in order to measure students' English proficiency (Institute for International Business Communication, 2015; Council on International Educational Exchange, 2013). Although the results of these standardized tests can be useful for school administrators and English teachers, administering them requires time and money; it takes a student two hours to complete each test, both of which cost more than 3,000 yen per person. Because of this there have been calls for a less expensive and more efficient alternative to these

standardized English proficiency tests. In an effort to address this need, Maki, Wasada and Hashimoto (2003) developed the Minimal English Test (MET).

The MET is a dictation test consisting of two passages with 72 blanks that the test-taker needs to fill in; the test-taker listens to pre-recorded monologues and writes down a word in each blank. Because of the gap-filling text completion format used by the MET it could be viewed as a type of a cloze test. Cloze tests were first introduced by Taylor (1953) as a measure of the reading ability of native English speakers. Cloze tests attracted a lot of attention in EFL literature from the 1960s to the 1980s as a means of measuring EFL learners' reading abilities or overall English proficiency (see Brown, 2013, for issues regarding cloze testing).

However, while similar in some respects, the MET has two distinct features that are different from most cloze tests; auditory cues are given to the test-taker and the number of words between one blank

---

Kanzaki M. (2016). Minimal English test versus the TOEIC. In G. Brooks, M. Porter, D. Fujimoto, & D. Tatsuki (Eds.), *The 2015 PanSIG Journal* (pp. 78-85). Tokyo, Japan: JALT.

to the next varies. This is because the MET has its roots not in the cloze test but in a Japanese language test for non-native speakers of Japanese called the Simple Performance-Oriented Test (SPOT), which was developed by Kobayashi, Ford and Yamashita (1995). The SPOT consists of 60 unrelated sentences, each of which has one purposefully-chosen *hiragana* character blanked out; the test-taker listens to a pre-recorded audio of the sentences and fills in the blanks. It takes only a few minutes to complete the SPOT. Kobayashi et al. (1995) reported a correlation coefficient of .82 between the scores of the SPOT and Tsukuba University's placement test, which consists of vocabulary, grammar, listening and reading sections and requires 150 minutes to complete. The MET was modeled after the SPOT, and Goto, Maki and Kasai (2010) called the MET "an English version of the SPOT" (p. 95).

Some correlation studies have been carried out to investigate the relationships between MET scores and scores of other English tests, such as the English test in the university entrance examination in Japan ( $r$ s between .60 and .72, reported in Goto et al., 2010), the TOEIC ( $r = .74$ , reported in Maki, Hasebe & Umezawa, 2010), the STPE Eiken 2nd Grade ( $r = .59$ , reported in Maki & Hasebe, 2013), and the Vocabulary Levels Test ( $r = .81$ , reported in Kasai, Maki & Niinuma, 2005). For a comprehensive list of papers on the MET, see Maki (2015). However, there has been no correlation study comparing the results of the MET and a speaking test. This study was therefore designed to compare the results of the MET and the TOEIC speaking test for correlations. This would then be compared to the correlation between the results of MET and the TOEIC listening and reading tests.

## Method

The MET and the listening, reading and speaking tests of the TOEIC (hereafter TC, TR and TS) were administered to 90 participants, and the scores of the four tests were then analyzed for correlations.

## Participants

The study participants were 90 students who

were attending a private university specializing in foreign languages in the Kanto area. They agreed to participate in the study in exchange for a cash reward of 1,000 yen, although they had to pay the 3,500 yen to take the TOEIC listening and reading tests on July 29, 2014. Initially, 94 students were to take part, but four of them were excluded because their MET scores were below 30 out of 72, which indicated that they had not taken the test seriously. The purposes of the study as well as the procedures and requirements were explained to the participants before they signed a consent form.

Among the 90 participants, seven were in their first academic year, 47 were in their second, 16 were in their third, and 20 were in their fourth; 15 of the participants were male and 75 were female. In terms of fields of study, there were 51 international communication majors, 26 English language majors, 11 international business majors, one Chinese language major and one Portuguese language major. All the participants were native Japanese speakers except for two native Korean speakers and one native Chinese speaker. One of them was enrolled in the TOEIC-860 course, seven in the TOEIC-730 courses, 42 in the TOEIC-650 courses and six in the TOEIC-600 courses (860, 730, 650 and 600 indicate the target scores of the courses). The remaining 34 were not taking any TOEIC course.

## Materials

The MET and the three types of the TOEIC (TL, TR and TS) were used in this study. The TL and TR are always administered together and are therefore usually treated as two sections of one test. The TS, on the other hand, can be taken independently when it is administered as part of the Institutional Program (IP). With the TOEIC IP the institution that is administering the test sets the time, date and place of the exam. The three TOEIC tests used in the study were administered as IP tests. These tests were not computerized adaptive tests and all of the participants in this study would have received the same questions in the three tests.

**Minimal English Test (MET).** The MET consists of two passages; one with 200 words and the other

with 198 words. Both of these passages are taken from an English textbook for university students written by Kawana and Walker (2002). The accompanying audio for this book is also used during the MET. The two passages are spread out over 36 lines of between six and 17 words and the average number of words per line is 11. Each line has two blanks, and only words that have four letters or fewer are blanked out. Because of this restriction, the deletion frequency of the MET is not regular; the number of words between two blanks is between 0 and 10, with an average word number of 4.24 between blanks, excluding the interval between the last blank of the first passage and the first blank of the second passage. For the actual test sheet along with the answer key, see the Appendix. The test-taker listens to the audio, which is set at a speed of 125 words per minute, and fills in the 72 blanks. There is a short pause of 10 seconds between the two passages (between lines 18 and 19). Because auditory cues are given, only the intended word is accepted as the correct answer, and semantically acceptable alternatives are counted as wrong as were spelling mistakes. However, in this study the author made one exception for the misspelling of “paid” in line 9, and

answers such as “payed” and “paied” were accepted on the ground that those who misspelled this were able to hear the word correctly and knew that it was the past form of “pay”.

**TOEIC Listening Test (TL).** The TL consists of 100 multiple-choice questions, and raw scores of between 0 and 100 are converted to scaled scores of between 5 and 495. The TL has four parts, the details of which are shown in Table 1.

**TOEIC Reading Test (TR).** The TR consists of 100 multiple-choice questions, and raw scores of between 0 and 100 are converted to scaled scores of between 5 and 495. The TR has three parts, the details of which are shown in Table 2.

*Note.* The TR starts with Part 5 because it immediately follows the TL, which ends with Part 4, and the two tests are always taken as a set.

**TOEIC Speaking Test (TS).** The TS is a computer-based test requiring the test-taker to sit in front of a computer wearing a headset with a microphone. Instructions are provided on the computer screen and through the headset. The test-taker speaks into the microphone and their speeches are recorded and sent to certified raters for evaluation.

TABLE 1

*Four Parts of the TL*

Part	Task	# of Qs
1	For each question with a photo, listen to four sentences and choose the one that best describes the image.	10
2	Listen to a question or statement followed by three responses and choose the most appropriate response.	30
3	Listen to a conversation and answer comprehension questions.	30
4	Listen to a short talk and answer comprehension questions.	30

TABLE 2

*Three Parts of the TR*

Part	Task	# of Qs
5	Choose a word or phrase to fill in a blank in a sentence.	40
6	Choose words or phrases to fill in blanks in a passage.	12
7	Read a passage or a set of two passages and answer comprehension questions.	48

There are 11 questions in the TS and scores are given in the range of 0 to 200. Table 3 shows the details of the TS.

## Results

All the statistical analyses were conducted using IBM SPSS Statistics for Windows version 22.

### Descriptive Statistics

Table 4 shows the descriptive statistics for the scores of the MET and the three TOEIC tests. The average TL score is 113.44 points higher than the average TR score. The average combined score of the TL and TR is 625.78.

### Reliability

The reliability index (Cronbach's alpha) for the MET

is .85, which is quite high for a simple test that can be completed in five minutes. The reliability indices of the TL, TR and TC for a particular test form are not available because the Educational Testing Service, which administers the TOEIC tests, does not disclose them. However, the Educational Testing Service (2013) reported that the reliability index (KR-20) of the TL and TR scores across all forms of their norming samples is "approximately .90" (p. 16). Also, the Educational Testing Service (2010) reported that the reliability of the TS is .80 "based on the data from January 2008 to December 2009 administrations in the Public Testing Program" (p. 18).

### Correlations

**Three TOEIC tests.** Table 5 shows the correlations between the scores of the three TOEIC tests. Among

TABLE 3  
*Details of the TS*

Question #	Task
1-2	Read aloud the text that appears on the screen.
3	Describe the picture on the screen.
4-6	Answer three questions about a single topic as though you are participating in a telephone interview.
7-9	Read the information on the screen and answer three questions about it as though you are responding to a telephone inquiry.
10	Listen to a recorded message about a problem and propose a solution for it.
11	Express an opinion about a specific topic.

TABLE 4  
*Descriptive Statistics for MET, TL, TR and TS (N = 90)*

Test (possible scores)	Mean	SD	Minimum	Maximum
MET (0-72)	47.49	8.55	31	67
TL (5-495)	369.61	56.67	170	495
TR (5-495)	256.17	71.52	100	430
TLR (10-990)	625.78	116.22	310	925
TS (0-200)	112.36	21.39	60	170

*Note.* TLR = TL and TR combined.

the three combinations, the highest is between the TL and TR scores ( $r = .64$ ) and the lowest is between the TR and TS scores ( $r = .38$ ). The correlation between the TL and TS comes between them ( $r = .45$ ). This order is in agreement with the correlations reported by Kanzaki (2015), which are .68 between the TL and TR, .50 between the TL and TS, and .48 between the TR and TS.

The correlation between the TS scores and the combined scores of the TL and TR is .45 ( $p < .001$ ).

**MET vs. TOEIC.** Table 6 show the correlations between the MET and the three TOEIC tests. The MET correlates with the TS at .59 and the figure is higher than those between the MET and the TL ( $r = .39$ ) and the MET and the TR ( $r = .51$ ).

## Discussion

One of the most striking aspects of the results is the MET scores correlate with the TS scores more strongly than with the TL scores or with the TR scores. The MET only requires the test-taker to listen and write down what is heard, and so it does not measure speaking skills. It is unclear how the MET, a gap-filling dictation test, can be said to evaluate the speaking abilities of learners. However, there are two features of the MET that might relate to speaking abilities. One is the multitasking nature of the MET. When taking

the MET, one has to listen to the audio, read the text, write down words and anticipate what will come next. Similar multitasking abilities are needed for speaking; when we speak with someone, we have to listen, speak and think about what to say next at the same time.

The other point is the processing speed needed to perform well on the MET. In order to complete the task satisfactorily on the MET, the test-taker has to process information quickly when listening and reading in order to write down what they hear. This type of quickness is also necessary for speaking, especially in the TS, where the test-taker has to come up with something to say within a certain amount of time and then has to say what he or she wants to say within a given timeframe.

Another striking aspect of the results is the correlation between the MET and TL, which is lower than those between the MET and TS and between the MET and TR. This is surprising, considering that the MET consists of auditory cues that the test-taker has to listen to and complete a task based on what they hear. Moreover, ordinary cloze tests, which do not provide auditory cues, "have consistently correlated best with measures of listening comprehension" (Oller, 1973, p. 114). It is puzzling that a test with listening elements correlates poorly with a listening test while the ones that do not have listening elements correlate well

TABLE 5  
*Correlations Between 3 TOEIC Tests (N = 90)*

	TL	TR	TS
TL	1	.64*	.45*
TR		1	.38*
TS			1

\* =  $p < .001$

TABLE 6  
*Correlations Between MET and TOEIC (N = 90)*

	TL	TR	TLR	TS
MET	.39*	.51*	.50*	.59*

Note. TLR = TL and TR combined.

\* =  $p < .001$



with a listening test. One possible explanation is that the MET does not function as a tool for measuring the listening comprehension of learners because the audio for the MET is not designed to test listening comprehension. It only gives auditory cues and what it tests is whether or not the test-taker recognizes words while the TOEIC listening test measures the test-taker's ability to listen to and process sentences of varying degrees of difficulty. Thus the test-taker's ability to multitask and move from one question to the next quickly plays a more important role than their listening ability on the MET. If you are not able to listen while writing down a word, you will miss the next word, and if you are slow to answer, it will be hard to catch up with the pace.

## Conclusion

The results suggest a possibility that the MET may be able to predict the test-taker's speaking abilities better than the TL and TR, as there was a stronger relationship between the TS and MET than between the TS and the TL and TR with the group of students involved in this study.

One limitation of the study is the composition of the participants; they were all language majors and most of them regularly had English lessons with native English-speaking teachers and they were more used to speaking English than the average English learner in Japan, and therefore the findings cannot be generalized. Another limitation is that a sample size of 90 participants is not large enough to make definitive generalizations about the MET. Finally, it should be remembered that the results presented above were obtained from a particular group of participants taking particular forms of the TL, TR and TS; if different participants take different forms of the tests, the results may not be the same.

## Acknowledgements

*The author would like to express his gratitude to Dr. Hideki Maki of Gifu University for providing the MET-related materials. Without his support, this study would not have been possible. Also, the author is grateful to Professor Norihito Kawana of Sapporo International*

*University and Seibido Shuppan Cooperation for granting permission to reproduce the MET in this paper. This study was supported by JSPS KAKENHI Grant Number 25370727.*

## References

- Brown, J. D. (2013). My twenty-five years of cloze testing research: So what? *International Journal of Language Studies*, 7(1), 1–32.
- Council on International Educational Exchange (2013). *TOEFL web magazine: Yuuza No Koe* [Voice from a user] [Report on the web]. Retrieved from <http://www.cieej.or.jp/toefl/webmagazine/educators-voice/1303/>
- Educational Testing Service. (2010). TOEIC user guide: Speaking and writing. Princeton, NJ: Educational Testing service. Retrieved from [http://www.ets.org/s/toEIC/pdf/toEIC\\_sw\\_score\\_user\\_guide.pdf](http://www.ets.org/s/toEIC/pdf/toEIC_sw_score_user_guide.pdf)
- Educational Testing Service. (2013). TOEIC user guide: Listening and reading. Princeton, NJ: Educational Testing service. Retrieved from [http://www.ets.org/Media/Tests/Test\\_of\\_English\\_for\\_International\\_Communication/TOEIC\\_User\\_Gd.pdf](http://www.ets.org/Media/Tests/Test_of_English_for_International_Communication/TOEIC_User_Gd.pdf)
- Goto, K., Maki, H., & Kasai, C. (2010). The Minimal English Test: A new method to measure English as a Second Language proficiency. *Evaluation & Research in Education*, 23(2), 91-104. doi: 10.1080/09500791003734670
- Institute for International Business Communication (2015). *TOEIC newsletter vol. 125* [Online newsletter]. Retrieved from [http://www.toEIC.or.jp/library/toEIC\\_data/sys/letter/TOEICNewsletter125.pdf](http://www.toEIC.or.jp/library/toEIC_data/sys/letter/TOEICNewsletter125.pdf)
- Kasai, C., Maki, H., & Niinuma, F. (2005). The Minimal English Test: A strong correlation with the Paul Nation Proficiency Test. *Bulletin of the Faculty of Regional Studies, Gifu University*, 17, 45–52. Retrieved from <https://repository.lib.gifu-u.ac.jp/bitstream/123456789/4588/1/KJ00004182420.pdf>
- Kawana, N., & Walker, S. (2002). *This is Media.com*. Tokyo, Japan: Seibido.

- Kobayashi, N., Ford, J., & Yamamoto, H. (1995). *Nihongo Nouryoku Kani Shiken (SPOT) No Tokuten Bunpu Keikou: Tyuukyuu Muke Tesuto To Syokyuu Muke Tesuto* [Distribution of scores in the Simple Performance-Oriented Test (SPOT): Comparison of scores between easy and difficult versions of SPOT]. *Journal for Japanese Language Education*, 10, 107-119.
- Kanzaki, M. (2015). TOEIC survey: Speaking vs. listening and reading. In P. Clements, A. Krause & H. Brown (Eds.), *JALT2014 Conference Proceedings* (pp. 626-636). Tokyo: JALT.
- Maki, H. (2015) The Minimal English Test (MET) [Webpage]. Retrieved from <http://www.geocities.jp/makibelfast/MET.html>
- Maki, H., & Hasebe, M. (2013). The Minimal English Test and the test in practical English proficiency by the STEP. *Bulletin of the Faculty of Regional Studies, Gifu University*, 32, 25-30. Retrieved from [http://repository.lib.gifu-u.ac.jp/bitstream/123456789/45802/1/reg\\_030032003.pdf](http://repository.lib.gifu-u.ac.jp/bitstream/123456789/45802/1/reg_030032003.pdf)
- Maki, H., Hasebe, M., & Umezawa, T. (2010). A Study of correlation between the scores on the Minimal English Test (MET) and the scores on the Test of English for International Communication (TOEIC). *Bulletin of the Faculty of Regional Studies, Gifu University*, 27, 53-63. Retrieved from [http://repository.lib.gifu-u.ac.jp/bitstream/123456789/34847/1/reg\\_030027005.pdf](http://repository.lib.gifu-u.ac.jp/bitstream/123456789/34847/1/reg_030027005.pdf)
- Maki, H., Wasada, H., & Hashimoto, E. (2003) *Saishoo Eego Tesuto: Shoki Kenkyuu* [The Minimal English Test: A Preliminary Study]. *Eego Kyooiku*, 53(10), 47-50.
- Oller, J. W., Jr. (1973) Cloze tests of second language proficiency and what they measure. *Language Learning*, 23(1), 105-118. DOI: 10.1111/j.1467-1770.1973.tb00100.x
- Taylor, W. L. (1953). Cloze procedure: A new tool for measuring readability. *Journalism Quarterly*, 30, 415-453.

---

### Author's Biography:

**Masaya Kanzaki** teaches at Kanda University of International Studies. His research interests include vocabulary acquisition, language testing and corpus linguistics. He can be reached at [kanzaki-m@kanda.kuis.ac.jp](mailto:kanzaki-m@kanda.kuis.ac.jp).

---

## Appendix

### MET with answer key

1. The majority of people have at least one pet at (some) time in their (life).
2. Sometimes the relationship between a pet (dog) or cat and its owner is (so) close
3. that (they) begin to resemble (each) other in their appearance and behavior.
4. On the other (hand), owners of unusual pets (such) as tigers or snakes
5. sometimes (have) to protect themselves (from) their own pets.
6. Thirty years (ago) the idea of an inanimate (pet) first arose.
7. This was the pet (rock), which became a craze (in) the United States and
8. spread (to) other countries as (well).
9. People (paid) large sums of money for ordinary rocks and assigned (them) names.
10. They tied a leash around the rock and pulled (it) down the street just (like) a dog.
11. The rock owners (even) talked (to) their pet rocks.
12. Now (that) we have entered the computer age, (we) have virtual pets.
13. The Japanese Tamagotchi—(the) imaginary chicken (egg)—
14. (was) the precursor of (many) virtual pets.
15. Now there (are) an ever-increasing number of such virtual (pets)
16. which mostly young people are adopting (as) their (own).
17. And (if) your virtual pet (dies),
18. you (can) reserve a permanent resting place (on) the Internet in a virtual pet cemetery.
19. Sports are big business. Whereas Babe Ruth, the (most) famous athlete of (his) day,
20. was well-known (for) earning as (much) as the President of the United States, the average
21. salary (of) today's professional baseball players is (ten) times that of the President.
22. (And) a handful of sports superstars earn 100 times (more) through their contracts
23. (with) manufacturers of clothing, (food), and sports equipment.
24. But every generation produces (one) or two legendary athletes (who) rewrite
25. the record books, and whose ability and achievements (are) remembered (for) generations.
26. (In) the current generation Tiger Woods and Michael Jordan are two (such) legendary
27. figures, (both) of whom (have) achieved almost mythical status.
28. The (fact) that a large number of professional athletes (earn) huge incomes
29. has (led) to increased competition throughout (the) sports world.
30. Parents (send) their children to sports training camps (at) an early age.
31. Such (kids) typically practice three to (four) hours a day,
32. (all) weekend (and) during their school vacations
33. in order (to) better their chances of eventually obtaining (a) well-paid position
34. on a professional (team) when they grow (up).
35. As for the (many) young aspirants who do (not) succeed,
36. one wonders if they (will) regret having (lost) their childhood.