



## How effective are TOEIC<sup>®</sup> preparation guidelines?

journal or publication title	Journal of Chikushi Jogakuen University and Junior College
number	6
page range	15-25
year	2011-01-31
URL	<a href="http://id.nii.ac.jp/1219/00000106/">http://id.nii.ac.jp/1219/00000106/</a>

# How effective are TOEIC<sup>®</sup> preparation guidelines?

David John WOOD

## 1: Introduction

### 1.1: *Background and Test-taking Strategies*

This research concerns Japanese examinees' use of what Educational Testing Services publish in candidate manuals as a preparation guide for the Test of English for International Communication. Test users such as teachers with students taking exams are as responsible for test practices as test producers. While ETS suggests specific preparation is impossible, one section of their handbook (Appendix A, available to candidates some time prior to the test) gives about the same number of suggestions for ELP and test-specific preparation. This study seeks to establish a relationship between test strategies and performance, extending the existing area of research on test-taking strategies by comparing how candidates who concentrate more on English language proficiency than test-specific preparation tend to be more successful. We expect a stronger positive correlation between test results and ELP-specific preparation because previous studies have indicated this likelihood. While it is reasonable to expect those candidates who follow ETS guidelines positively would score more highly than those who do not, if there is a significant difference between ELP and test-specific approaches, this may reveal a disparity.

### 1.2: *Research Question and Hypothesis*

The central research question that we investigate is: "What difference if any is there between the effectiveness of the subjects' frequencies of use of ETS ELP-specific preparation and test-specific approaches in terms of their results?"

The hypothesis is as follows: "There is a significant efficient relationship between factor one, the frequency of test preparation strategies' use, and factor two, their TOEIC results."

## 2: Literature Review

### 2.1: *Japanese Background*

Introduced in 1979 after a request by Japan's MITI, the popularity of TOEIC has expanded dramatically to 5,000,000 takers, including 80% in Japan and Korea. It is "widely held to be the test for assessing business English skills" (Brown, 2006, p. 1177). However, Japanese candidates' average is lower than most nationalities, averaging 451 (Stoynoff & Chapell, 2005) partly because of the strong belief that test-specific preparation is efficient.

The focus has shifted to an international multi-level testing educational instrument. In efficiency-minded Japan, employment orientation starts early, extending to data collection by careers offices. The solid face validity of TOEIC, along with consumer demand and producer marketing, has contributed to the transformation, accounting for the urgency to score higher.

## **2.2: Test-taking strategies**

However, the best way to prepare effectively remains elusive. The development of test-taking strategies has considerable relevance in this context. While testing has a long history, understanding what test-takers actually do in tests and how behaviors match the abilities that test writers try to test is a recent endeavor (Cohen, 2006). The relationship between learner characteristics like ELP level and strategies has come more into focus. According to Cohen, weaker test-takers compensate for lack of proficiency by concentrating more on test-specific preparation.

As regards strategy instruction for performance on high-stakes standardized tests, Forster and Karn (1998) made a detailed study of both test-specific and more general ELP strategies in Japan. There is a degree of overlap plus areas of contention. Areas of overlap include: understanding directions before the test saving time during it; working quickly by staying ahead of the recording in the listening section; and, not spending too much time on any particular question.

Contentious strategies include guessing systems when time is short to complete unanswered questions, resembling ETS implications that guessing is expected. Culturally this may be dubious. The most controversial point of all is: "Studying for (TOEIC) will dramatically increase English proficiency" (Forster & Karn, 1998, p. 14). TOEIC does not cause English proficiency. It only calculates certain aspects of it.

One study that seems to indicate test-specific strategies are at best an inferior and makeshift alternative to ELP-specific preparation has been that made by Tian (cited in Cohen, 2006) in reference to the TOEIC academic twin, TOEFL. Tian suggests that high scorers employed ELP foremost in understanding passages, and used test-specific strategies taught them only in an auxiliary fashion. Candidates getting poorer results tend to overuse test-specific strategies in favor of striving to comprehend on the basis of their proficiency. Tian's research accordingly "serves as a warning that (test-specific) strategy training materials may not necessarily help those who need it the most (those with lower ELP) and perhaps most benefit those who least need assistance (those with higher ELP)" (cited in Cohen, 2006, p. 324).

## **2.3: Ethicality versus Purpose**

Teacher attitudes to ELP exam preparation are mixed. Pressure has eroded resistance but concern underpins perceived benefits. Concluding findings about extensive direct test preparation, Robb and Ercanbrack commented: "Forcing students to study TOEIC preparatory material might, therefore, be doing them a disservice if communicative ability is the goal of the program" (1999, p. 18). "Disservice" here means subverting the purpose of the exam to test communication and frustrating candidates who may risk being misdirected

if they sacrifice long-term ELP improvement for short-term “test-wiseness” increase (Cohen, 2006).

Objections about test preparation ethicality are long-standing. Hamp-Lyons’ bemoaned an “enormous (test) preparation industry” jeopardizing properly designed curricula, concluding: “It is a problem in program administration, in teaching, in textbook authoring, and in the educational-commercial interface in the TESOL profession” (1998, p. 335 and p. 336). Such doubt persists: “It has been suggested that a code of practice for the TESOL profession should address the issue of ethical preparation material and practices (which) seems justified in the light of the tension between market interests on the one hand and educational interests on the other” (Knapman, 2008, p. 92). In terms of the test alone, coaching students to score above what they might otherwise achieve breaches its stated aim (Schmidt, 2003).

An example of dubious purpose was reported at Asia University, Japan (Koelbleitner, Gustavsen & Alberding, 2003). Administrators pushed TOEIC for placement and exit uses stressing employers’ expected minimum 600 points hiring criterion for jobs involving English, plus two-hour orientation pre-test sessions. However, results were below target even for classes of higher ability students, contradicting existing placement levels and reflecting undue influence from the test’s high face value. The example indicates test-specific preparation potential counter-productivity and ignores ELP preparation as the appropriate approach. The test’s purpose cannot meaningfully be redirected for instant results.

#### ***2.4: Shortcomings of Direct Preparation***

Brown (2006) expressed reservation to direct test preparation as commercially available materials merely tend to simplify test item practice to the point where they do nothing to help lower-level learners handle real exam complexities, leading to frustration and even discontinuation of studies. Developing experience and interest in reading to increase comprehension is preferable. Direct preparation materials can have a negative effect as they focus on isolated point teaching, leading candidates to view the test as puzzle-solving instead of focusing on reading for meaning. Candidates develop low tolerance for ambiguity, failing to develop the ability to read listening questions quickly enough to answer.

#### ***2.5: Validity of TOEIC and Teachers’ Roles***

The newly recognized status of TOEIC is indicated by two studies providing perspectives for teachers, Stoyhoff and Chapelle (2005) and (Stoyhoff, 2009). A principle concern is the divide between ELP testers and teachers caused by the specialization of the former. Compounding the situation are changing definitions of validity dictating measurement theory. Originally, validity was considered a test characteristic (how far it measures what it was supposed to) but this is now “the extent to which test uses and interpretations can be justified” (Stoyhoff and Chapelle, 2005, p. 6). The concept of reliability has become another type of validity evidence. Instead of evaluating a test by asking if it possesses one of the original kinds of validity, we can consider its validity for our specific use, tantamount to context validity. Defining test-use validity is the responsibility of the teacher not test writer, increasing the importance of teacher-generated washback.

The polarization of testing and assessment must be broken down by understanding such factors like justification. To achieve teacher/learner responsibility for a test, careful examination of the handbook is vital. The handbook bears directly on this present study as it includes preparation suggestions for candidates. Research devoted to ETS guidelines is very limited with no relevant publications. Investigating the effectiveness of test preparation strategies is therefore an important starting point. While simultaneously developing classes, courses and conditions dealing directly with score improvement, teachers can regain the initiative by grasping the overall implications. Stoyhoff and Chappelle (2005) point to this (Chapter 3: “Using the Test Manual to Learn More About a Test”) but stop short at pinpointing the need to identify which abilities or skills a test is designed to assess, without proceeding to issues like test preparation guidelines.

## **2.6: Conclusion**

Preparation is perceived by candidates and validated by research as potentially useful, but which kind is efficient is crucial. The research discussed suggests test-specific preparation is neither helpful nor valid, despite its considerable contingency of proponents, especially in Japan. ELP preparation is true to the test’s international communication goals and most users. Redirecting the preparation process risks defeating the original purpose. Because of shared responsibility users and producers need to ensure the test comprehensively contributes to its true purpose, including the handbook. This pilot study indicates potential further studies.

## **3: Methodology**

### **3.1: Participants**

Subjects were a homogeneous group of recent Japanese university graduates who had taken TOEIC at the same time under the same conditions. They had taken no test-specific classes as their English department had no TOEIC program. Apart from one high achiever, all subjects belonged to the same academic grouping, namely one rank above the school’s average.

### **3.2: Materials**

A questionnaire was devised to gather information relevant to the research question, nine questions each dealing with test-specific and ELP-specific preparation. As preliminary background, the examinee handbook section “How To Get Ready To Take the TOEIC Test” (Appendix A) was closely studied. Its points became central to the questionnaire with nine questions about each. The questionnaire was constructed to establish: the homogeneity of the group in terms of gender, university major, year of taking the test, and so on; determine their TOEIC result and academic ability confidentially; and, determine the frequency of both kinds of approach.

### 3.3: Procedures and Details of Design

The questionnaire was administered to one out of ten senior university English Department seminars during the year after they had graduated. The subjects' approaches to taking the exam were examined in direct reference to the officially recommended preparation guidelines. This information was correlated to their TOEIC results.

All communication was conducted via e-mail as the subjects were working in various places around Japan. This made communication difficult as most were busy and seldom used e-mail, but enhanced the independence of each subject's responses.

In the questionnaire, subjects' TOEIC results were the independent variable and the relative effectiveness of the strategies was operationalized by the frequency of test takers' application of the strategies and the relative correlation with their TOEIC results. A Likert-scaled multiple-choice format allowed subjects a reasonably unbiased range of answering options.

### 3.4: Analysis

The data collected from the questionnaire are shown in Tables 1, 2 and 3 below.

**Table 1: Subjects' ELP-specific Preparation Frequencies total number of responses plus subject numbers in parentheses in sequential order of TOEIC scores (see Table 3)**

Preparation Item:	Frequency:	Never	Some-Times	Usually	Often	Always
(i) "I immersed myself in English."			<b>3</b> (1,2,4)	<b>3</b> (3,5,6)	<b>4</b> (7,8,9,10)	
(ii) "I listened to recordings in English."			<b>4</b> (1,2,3,4)	<b>4</b> (5,6,7,9)	<b>2</b> (8,10)	
(iii) "I studied business English."		<b>4</b> (1,2,3,9)	<b>6</b> (4,5,6,7,8,10)			
(iv) "I spoke in English with friends."		<b>1</b> (2)	<b>9</b> (All others)			
(v) "I took related English courses."			<b>3</b> (1,2,4)	<b>4</b> (3,5,6,7)	<b>3</b> (8,9,10)	
(vi) "I used official TOEIC preparation materials."		<b>2</b> (1,2)	<b>3</b> (3,4,9)	<b>1</b> (10)	<b>3</b> (6,7,8)	<b>1</b> (5)
(vii) "I watched TV programs in English"			<b>2</b> (1,4)	<b>1</b> (2)	<b>7</b> (3,5,6,7,8,9,10)	
(viii) "I watched videos in English."			<b>3</b> (1,2,3)	<b>2</b> (4,6)	<b>5</b> (5,7,8,9,10)	
(ix) "I practiced by reading things written in English."			<b>4</b> (1,2,3,4)	<b>3</b> (6,7,8)	<b>3</b> (5,9,10)	

**Table 2: Subjects' Test-specific Preparation Frequencies total number of responses plus subject numbers in parentheses in their sequential order of TOEIC results (see Table 3)**

Preparation Item:	Frequency:	Never	Some-Times	Usually	Often	Always
(i) "I tried to answer questions to the best of my ability."				4 (2,3, 8,10)	6 (1,4,5, 6,7,9)	
(ii) "I answered questions quickly."			2 (1,5)	6 (2,4,6, 7,8,10)	2 (3,9)	
(iii) "I paid attention to the time during the Reading section."			1 (5)	4 (6,7, 8,10)	3 (1,2,4)	2 (3,9)
(iv) "I returned to questions I could not answer at first."	1 (3)		4 (5,6,7,8)	1 (2)	4 (1,4, 9,10)	
(v) "I studied the test format before the exam."			5 (1,2,3, 5,6)	1 (7)	4 (4,8, 9,10)	
(vi) "I studied the sample questions in the handbook."	3 (1,2,3)		1 (8)	1 (10)	4 (4,6,7,9)	1 (5)
(vii) "I visited the TOEIC website."	2 (1,2)		4 (3,4,5,10)	1 (6)	3 (7,8,9)	
(viii) "I guessed answers."			4 (1,2,8,9)		6 (3,4,5, 6,7,10)	
(ix) "I worked carefully during the test."				5 (5,6,7, 8,10)	4 (1,2,4,9)	1 (3)

**Table 3: Subjects' TOEIC R results**

Subject	1	2	3	4	5	6	7	8	9	10
Result	595	610	625	650	655	690	715	750	775	790

## 4: Data Analysis

### 4.1: Initial Analysis

The relative overall effectiveness of language and test-specific suggestions from the ETS examinee handbook was operationalized by the frequency of the subjects' application of the test preparation strategies and the relative correlation with their TOEIC results. To measure the relative effectiveness of the test-taking strategies a Pearson's correlation test was run to test the correlation coefficient between the frequency of application of test-taking strategies and the subjects' TOEIC scores. Subjects' responses were recorded on a 5-point Likert scale. Findings are in Table 4.

A Pearson's correlation analysis yielded a main effect for the effectiveness of TOEIC preparation strategies,  $R = .355$  ( $p < .001$ ). This allows further examination of the results to be conducted below in the form of splitting the data sets into test-specific and ELP-specific preparation strategies.

### 4.2: Medial Analysis

On the basis of the significant correlation determined in the initial analysis of the data above, further

**Table 4: Correlation of overall relationship between preparation frequency and TOEIC results**

Descriptive Statistics - Condition of testing = Frequency; Type of testing = ELP and test-specific preparation

	Mean	Standard Deviation	N
SCORE	2.9611	1.07985	180
TOEIC	685.5000	66.30744	180

Correlations - Condition of testing = Frequency; Type of testing = ELP and test-specific preparation

		SCORE	TOEIC
SCORE	Pearson Correlation	1	.355(**)
	Significance (2-tailed)	.	.000
	N	180	180
TOEIC	Pearson Correlation	.355(**)	1
	Significance (2-tailed)	.000	.
	N	180	180

\*\* Correlation is significant at the 0.01 level.

**Table 5: Correlation analysis for test-specific preparation frequency and TOEIC results**

Descriptive Statistics - Condition of testing = Frequency; Type of testing = Test-specific preparation

	Mean	Standard Deviation	N
SCORE	4.4000	.59587	90
TOEIC	685.5000	66.49343	90

Correlations - Condition of testing = Frequency; Type of testing = Test-specific preparation

		SCORE	TOEIC
SCORE	Pearson Correlation	1	.185
	Sig. (2-tailed)		.081
	N	90	90
TOEIC	Pearson Correlation	.185	1
	Sig. (2-tailed)	.081	.
	N	90	90

analysis of the test results is justified. Two types of preparation strategies have been identified for further investigation: test-specific preparation strategies and English language proficiency specific strategies in Tables 5 and 6 below.

A Pearson's correlation analysis yielded a main effect for the effectiveness of TOEIC ELP-specific preparation strategies,  $R = .537$  ( $p < .001$ ). The analysis of the test-specific preparation strategies did not yield any specific results ( $R = .185$   $P > .05$ ). Accordingly, individual ELP-specific preparation points are investigated further below.

### 4.3: Final Analysis

The significant correlation between ELP-specific preparation and TOEIC results allows us to further ana-

**Table 6: Correlation analysis for ELP-specific preparation frequency and TOEIC results**

Descriptive Statistics - Condition of testing = Frequency; Type of testing = ELP-specific preparation

	Mean	Standard Deviation	N
SCORE	2.7556	1.00907	90
TOEIC	685.5000	66.49343	90

Correlations - Condition of testing = Frequency; Type of testing = ELP-specific preparation

		SCORE	TOEIC
SCORE	Pearson Correlation	1	.537(**)
	Sig. (2-tailed)		.000
	N	90	90
TOEIC	Pearson Correlation	.537(**)	1
	Sig. (2-tailed)	.000	.
	N	90	90

\*\* Correlation is significant at the 0.01 level.

lyze the data set for the individual items. Results of the analysis are shown in Table 7.

As shown in Table 7 the correlation analysis for the relationship between subjects taking related English courses and their TOEIC results yielded significant results for 5 of the 9 items. The trend for all of these 5 items displayed strong correlation coefficients ranging from  $R = .755$  to  $R = .898$ , in every case with  $p < .001$ .

Item (v), taking related courses, yielded the strongest result of all the five significant results,  $R = .898$ , followed closely by the result for subjects immersing themselves in English, item (i),  $R = .891$ . The next

**Table 7: Results of the Pearson's correlation and analysis between the frequency of application of individual items of ELP specific preparation strategies and TOEIC results**

Item Type	Item	N	Mean	R	Significance
ELP-specific preparation(i)	"I immersed myself in English."	10	3.1	.891	** significant at the 0.01 level
ELP-specific preparation(ii)	"I listened to recordings in English."	10	2.8	.871	** significant at the 0.01 level
ELP-specific preparation(iii)	"I studied business English."	10	1.6	.423	
ELP-specific preparation(iv)	"I spoke in English with friends."	10	1.9	.381	
ELP-specific preparation(v)	"I took related English courses."	10	3.0	.898	** significant at the 0.01 level
ELP-specific preparation(vi)	"I used official TOEIC preparation materials."	10	2.8	.429	
ELP-specific preparation(vii)	"I watched TV programs in English"	10	3.5	.614	
ELP-specific preparation(viii)	"I watched videos in English."	10	3.2	.840	** significant at the 0.01 level
ELP-specific preparation (ix)	"I practiced by reading things written in English."	10	2.9	.775	** significant at the 0.01 level

strongest result yielded was for subjects listening to recordings in English, item (ii),  $R = .871$ . This was followed by the result yielded for subjects watching videos in English, item (viii)  $R = .840$ , and finally, the result that was yielded for students practicing by reading things written in English, item (ix),  $R = .775$ .

#### **4.4: Data Result Discussion**

The annual population of Japanese female TOEIC takers is approximately 1,000,000 so the limited sample in this study only questionably confirms the hypothesis, and whether subjects who got higher TOEIC results did so because they used ELP-specific approaches more frequently or otherwise is an irresolvable issue of directionality.

In line with studies that find ELP-specific strategies more efficient, lower scoring subjects' less frequent choice of ELP-specific strategies in this present study supports Cohen's findings that weaker test-takers try to compensate for their lack of proficiency by defaulting to test-specific strategies (2006, p. 312). Cohen found that one of the major themes in studies on test-taking strategies was the influence of candidates' ELP abilities on the frequency of the strategies that they employ. Lower ability candidates tend to try making up for this by concentrating more on how to take the test than on studying English in various ways. The findings of this study concur with Tian's conclusions (cited in Cohen, 2006) that high scorers employ ELP strategies foremost in understanding passages, and use the test-specific strategies they are taught in an auxiliary fashion, indicating inability to answer questions.

The central research question can be tentatively answered in terms of the results above. For this sample group there is a positive correlation between ETS ELP-specific preparation frequencies of use and their TOEIC result. The significant ELP items are generally the most available study pursuits for students, like English courses, as they were in an English department, whereas there was limited availability of business English and TOEIC materials. Immersing themselves in reading and listening to English indicates both motivation and application. Japanese students seldom speak English, perhaps out of lack of confidence. While TOEIC does not test speaking directly, it is still an approach worth encouraging. Watching videos is also of significance, partly explainable because many teachers use videos in class and students enjoy them. Clearly the belief that test-specific preparation is efficient (Forster and Karn, 1998) is brought into question by the results of this study.

## **5: Conclusion**

### **5.1: General Conclusion and Implications for Future Research**

ETS guidelines were determined as having a significant correlation overall with the subjects' TOEIC results. Further investigation determined that only ELP-specific preparation was significant, justifying a closer look at individual items. The final analysis demonstrated certain items' significance within the limitations of this sample group. Opinion about test-taking strategies was divided, though there was a general consensus

that preparation was important. Cohen (2006) indicated that ELP-specific was sounder. As ETS itself suggests test-specific preparation is impossible, it is confusing why their handbook mixes both kinds. No studies relating to TOEIC preparation contain investigations like the current study, so it may promote further consideration. Ideally ETS would also consider the implications for its handbook which examinees rely on each year. The correlation between test scores and general ELP-specific preparation necessitates further studies. Some of ETS's budget could be reinvested to determine how best to rewrite their guidelines accordingly. ETS should omit test-specific preparation guidelines until research justifies reinstatement. This could reform the strong Japanese contingent still promoting test-specific preparation, with countless new test-specific preparation texts appearing yearly.

### References

- Brown, H. 2006. Learner Perceptions of TOEIC Results & Language Skill Improvements. *JALT 2005 Conference Proceedings*, Tokyo, 1176-1181.
- Chapelle, C., & Stoyanoff, S. 2005. *ESOL Tests & Testing*. Virginia: TESOL Inc.
- Cohen, D. 2006. The Coming of Age of Research on Test-taking Strategies. *Language Assessment Quarterly*, 3.4, 307-331.
- ETS. 2008. *TOEIC Listening & Reading Test Examinee Handbook*. New Jersey: Princeton.
- Forster, D., & Kam, R. 1998. *Teaching TOEIC/TOEFL Test-taking Strategies*. Japan Women's University, Tokyo.
- Hamp-Lyons, L. 1998. Ethical Test Preparation Practice. *TESOL Quarterly* 32.2, 329-337.
- Knapman, G. 2008. The TOEIC: Critical Review. *Fukuikogyodai Journal* 38, 85-94.
- Koelbleitner, C. Gustavsen, E., & Alberding, M., 2003. An Examination of the Proposed Use of the TOEIC at Asia University. Tokyo: *CELE Journal* 11, 115-124.
- Richards, R. 1992. Review of TOEIC. In J. Conoley (Ed.) *The Eleventh Mental Measurements Yearbook*, University of Nebraska Press.
- Robb, T., & Ercanbrack, J. 1999. *A Study of the Effect of Direct Test Preparation on the TOEIC Scores of Japanese University Students*. [www.kyoto-su.ac.jp/information/tesl-ej/ej12/a2.html](http://www.kyoto-su.ac.jp/information/tesl-ej/ej12/a2.html). Retrieved October 9, 2009.
- Stoyanoff, S. 2009. Recent Developments in Language Assessment & the Case of Four Large-scale Tests of ESOL Ability. *Language Teaching* 42:1, 1-40 \_c. Cambridge: Cambridge University Press S0261444808005399.

### Appendix A

*TOEIC Examinee handbook 2008, Page 4 (Copyright ETS)*

“How To Get Ready To Take the TOEIC Test

The TOEIC test is not based on the content of any particular English course but rather on your English-language proficiency your overall ability to use English. Improvement in proficiency may take some time and is generally achieved through a combination of practice and study. The TOEIC test does not test business knowledge, and you are not required to know specialized business and technical vocabulary beyond what is used in everyday work activities.

Before taking the TOEIC test, there are several things you can do to prepare for the test and improve your English proficiency:

Become familiar with the test format and know how to mark your answers on the answer sheet. You can then focus your attention on the test questions themselves.

Carefully review the test directions and the sample questions on pages 8 - 13 and the sample answer sheet and sample Background Questionnaire on pages 15 - 18.

Immerse yourself in the language as frequently as possible and in as many ways as possible if it has been some time since you have had contact with English.

Reading, watching TV and videos, listening to recordings, taking an English course, and speaking with friends and colleagues are some of the ways to practice English.

#### Web Resource Guide

The official TOEIC website has many resources you may find helpful to familiarize yourself with the test:

- test preparation material
- scoring information
- frequently asked questions (FAQs)

#### During the Test

Work quickly and carefully.

Do not spend too much time on any one question.

Mark your answers on your answer sheet and not in the test book.

Mark only one answer for each question. If you mark more than one answer, that question will be counted wrong even if one of the answers you marked is correct.

You will receive credit only for answers marked in the circles on the answer sheet. Your score will be based on the number of questions you answer correctly. There is no penalty for guessing.

Try to answer every question to the best of your ability.

Pay close attention to the time during the Reading section of the test. In the Reading section (75 minutes) you have to pace yourself, so work quickly and if you do not know the answer to a question, come back to it later.

You may not use note paper.”

(デイビッド ジョン ウッド : 英語学科 教授)