

THE EFFECT OF TEST PREPARATION TOEFL READING TESTS

A.Elvi W. Simanjuntak
English Education Department
Faculty of Education and Teachers Training
University of Lancang Kuning
agnesiaelviwisnita@fkip.unilak.ac.id

Abstract

The main objective of this study was to see whether preparation influences the learners' reading comprehension of TOEFL? Also as a subsidiary question, this study tried to find answer(s) to the following question: Does getting a high score imply enhancement in the reading comprehension ability? To this end, sixty EFL learners were selected to take part in a preparation program in which they were instructed different techniques and strategies to deal with the reading section of TOEFL tests. Although the participants' scores enhanced in the post tests. Generally, findings showed that the preparation effect was more visible on performance of the TOEFL reading test. The TOEFL preparation group on the TOEFL reading test.

Key words: TOEFL, test preparation, reading

1. INTRODUCTION

“TOEFL or Test of English as a Foreign Language is an English Proficiency test, conducted to evaluate the English speaking and understanding ability of a candidate. TOEFL is accepted as proof of English proficiency in 9,000 colleges, universities and organizations and other TOEFL participating institutes across 130 countries in the World. TOEFL is conducted more than 50 times a year by Educational Testing Service (ETS)” (a US-based non-profit organisation). With the ever-increasing importance of TOEFL test for non-native speakers, as decisive gateways to better education opportunities and obtaining survival skills in the countries where English is spoken as first language, getting focused preparation for these instrumental tests has increasingly become a vital requirement. This special preparation may be of different sorts. For example, test familiarization is designed to ensure that prospective test takers are well versed in the general skills required for test taking and to help them gain familiarity with the procedures that are required to take a particular test. This type of preparation may entail, for instance, exposing test takers to the kinds of item formats they will encounter, making certain that they know when to guess, and helping them learn to apportion their time appropriately. Special preparation of this sort is generally regarded as desirable, as it presumably enables individuals to master the mechanics of test taking, thereby freeing them to focus on, and accurately demonstrate, the skills and abilities that are being assessed (Powers, 2012). A variety of different types of test preparation: formal commercial coaching, school-offered test preparation programs, and test sponsor-provided test familiarization have been introduced (Powers, 2012).

Test developers, researchers, teachers and students/test taker have always been concerned with the issue of the effect of test preparation on the learners' performance of TOEFL test. However, on the other hand, from a more theoretical and pedagogical point of view, researchers and test designers often highlight the merit of a test which can independently measure the underlying ability of test-takers regardless of how well they have prepared or crammed for the test. In other words, if preparing results in a higher test score regardless of whether the candidate's ability has increased, thus the construct validity of the test is questionable (Nguyen, 2007). From a practical stand point, however, teachers and students/test takers often feel the vitality of test preparation before being tested.

Given the importance attached to it, this skill was chosen as the independent variable for this study. Their expectations, however, are not always met. The reasons for the gap between their expectations and what actually occurs have not been researched in-depth to date in Iran. It is thus argued that from theoretical and practical outlooks, it is useful to investigate the effect of test preparation on test scores in general, and on reading test scores in particular in an Iranian setting. In this study, test candidates were instructed various combinations of test preparation materials to enable them

to become more familiar with the reading section of the TOEFL, with each of the item types they contain, and the strategies necessary to tackle this section in each test. The materials included full-length sample reading tests, and hints or tips for approaching each of these item types.

2. REVIEW OF RELATED LITERATURE

a. The Relationship between Test Preparation Programs and Test Score Improvement

Messick (1982, as cited in Powers, 2012) provided an insightful logical analysis of the ways in which special test preparation may impact validity, there appears to have been little empirical research to demonstrate how such practices may affect, for example, the relationship of test scores to other relevant measures. Most of the studies of test preparation have focused on the extent to which these practices cause spurious test score improvement. However, although relatively rare, researchers have also examined, in both a logical and an empirical manner, the effects of test preparation on the empirical relationships of test scores to other indicators of developed ability.

The impact of special preparation on test validity is a germane consideration. Though the assumption is sometimes made that preparation can serve only to adulterate the construct validity and impair the predictive power of a test, some kinds of special preparation may, by reducing irrelevant sources of test difficulty, actually improve both construct validity and predictive validity (Hayes & Read, 2004; Farnsworth, 2013). The findings from this body of research, however, are to some extent contradictory: while a relationship between test preparation and test performance is acknowledged in some studies, doubts about this relationship have also been voiced.

b. Research on the Effect of Preparation Programs on Language Tests

The effect of test preparation on language tests is a matter of some controversy; the effectiveness of preparation programs and products, unequal access to them, and concerns about the impact of preparation on the validity of test scores have always been an issue.

Farnsworth (2013) states that most language assessment research on test coaching has looked at the effect of test formats on instructional practices, or so-called test washback, and not looked at score increases due to coaching practice. Existing research, however, has shown similarly modest results of coaching on scores (Hamp-Lyons (1991); Brown, 1998; Nguyen, 2007). This effect is intrinsically difficult to measure however, because genuine proficiency gains and score gains derived from mastering the test format and/or test-taking strategies are very hard to tease apart, and real proficiency gains likely occur alongside “test wiseness” score gains during test coaching instruction. The issue that finds head from these studies on the effects of preparation programs is that researchers have been

at odds in their findings regarding the effect of language test preparation on test performance.

3. METHODOLOGY

• Research Contexts and Participants

Participants for the study were 30 students from 6th semester who were doing either the TOEFL preparation courses at Language Center in University of Lancang Kuning for the purpose of maximizing their chances of improving their reading score in TOEFL.

4. INSTRUMENTS

It was not possible to obtain operational versions of either test because of security consideration; therefore, the test materials used in the study were the TOEFL sample reading test versions. The number of items for each sample test was 50 questions and the time allotted was 60 minutes for each test administration.

5. DATA COLLECTION

All participants in the study took both the TOEFL practice tests. In order to avoid any possibility of a test practice effect, 30 participants took TOEFL reading test first. The data collection design can be summarized as follows:

Table 1. TOEFL Preparation Group

GROUP I (CLASS A)	GROUP II (CLASS B)
DN	JA
TR	FA
SL	FD
DS	EM
SR	DO
SD	DA
S	N
DF	ZM
NF	JS
FS	SE
SS	IC
YV	W
HP	RY
R	AA
FM	LW

6. DATA ANALYSIS

The effect of test preparation on test performance was investigated through a comparison of test performance. T-tests were employed for the mean score comparison. It should be noted that the finding in this study ought to be interpreted with caution since the number of test takers is admittedly small for

making generalizations, as maintained by (Camilli & Shepard, 1994; McNamara & Roever, 2006).

7. RESULTS

• The Effect of Preparation on the Performance of TOEFL Reading Test

There were less obvious differences in the TOEFL reading test scores across the two groups, as shown in table 2. The maximum and minimum scores of the two groups of test-takers are relatively close but low. Besides, the mean score and SD of the two groups, though different, were also close but low.

Table 2. Results of the TOEFL Reading Test

NO	STUDENTS'S INITIALS	SCORE			FINAL SCORES
		LISTENING	STRUCTURE	READING	
1	DN	43	36	42	403
2	TR	43	42	36	403
3	SL	43	35	42	400
4	DS	48	37	40	417
5	SR	52	45	46	417
6	SD	46	45	41	440
7	S	56	49	48	510
8	DF	57	53	47	523
9	NF	55	50	46	503
10	FS	44	43	45	440
11	SS	45	40	43	427
12	YV	61	61	57	597
13	HP	42	36	38	387
14	R	40	38	39	390
15	FM	36	39	37	373
16	JA	54	45	42	470
17	FA	47	43	39	430
18	FD	40	35	39	380
19	EM	44	66	57	557
20	DO	47	46	41	447
21	DA	47	41	41	430
22	N	50	37	31	393
23	ZM	43	39	32	380
24	JS	37	33	47	413
25	SE	42	44	38	510
26	IC	47	63	43	393
27	W	41	37	40	380
28	RY	43	38	33	403
29	AA	42	38	42	360
30	LW	37	34	37	

Nb: *The red words for class A and the black words for class b*

8. DISCUSSION AND CONCLUSION

The descriptive analyses of the raw scores, mean scores, and standard deviation of the two groups are suggestive of the effect of test preparation on test performance. Moreover, our findings are consistent with what Bachman et al. (1995) who claimed no effect of the TOEFL preparation on the learners' performance on the reading section. Likewise, in this study the results show no clear effect of the TOEFL reading preparation on the learners' performance on the TOEFL test. On the contrary, this study gainsays the findings reported by other researchers (e.g. Geranpayeh, 1994) interested in the same issue. Geranpayeh (1994) compared the effect of TOEFL preparation treatment with no preparation effect on the TOEFL test. He found that those who underwent TOEFL preparation outperformed those who received no instruction. Geranpayeh's study, however, is different from the present study in that it included an intact group to examine the effect of the independent variable, that is, the treatment. Another striking difference was the TOEFL test employed in Geranpayeh's (1994) study (Noguyen, 2007). In that study, the test had less construct validity than its subsequent versions, since it was more 'amenable' to test preparation than the more recent ones (Noguyen, 2007). We believe the results obtained in the TOEFL in this study were not affected by the test-taking strategies presented in the treatment.

On the other hand, in the TOEFL there are a set of limited question types available that can be answered not necessarily by drawing on the strategies needed. We also agree with Noguyen (2007) claim that combining different task types under one section would decrease the test method effect on the test performance, but it may have negative effect, instead. A third reason is the genre of the texts utilized in these two tests. One important issue noted by Miyasaka (2000) and Rubenstein (2004) is how test preparation can potentially affect test validity. Miyasaka (2000) stated that the majority of large-scale assessment tests should be designed so that one could make reasonable inferences about the achievement levels of students with respect to content knowledge and/or skills within a given domain. The primary concern is whether a few months of preparation can significantly influence scores. In other words, if preparation does affect test scores, is the purpose of the test to measure the knowledge of the test-taker (pertaining to the topic of the given test)? If yes, how much knowledge is truly being measured, specifically if test-takers have access to tools that may artificially enhance their scores (Rubenstein, 2004, p. 398)? ACT has noted that the earning of high scores on the ACT should not be merely a reflection of intrinsic talent or provisional preparation, but should reflect a level of accomplishment resulting from hard work, planning, and a solid commitment (ACT, 2005).

But the validity of the test score is compromised when test preparation artificially increases students' test scores without increasing mastery of the content domain, underlying subject-area knowledge, and/or testing skill (Perlman, 2004). Further, any test preparation practice that violates ethical

standards would in effect nullify any validity of the students' test results (Miyasaka, 2000). Cho and Bridgeman (2012) express concern about the predictive validity of the TOEFL expressed in terms of correlation with academic success. They maintain the students with higher TOEFL scores tended to earn higher grade point averages (GPA) in their study and the TOEFL provided information about the future academic performance of non-native English speaking students beyond that provided by other admissions tests. Combined with our results in this study, these observations can hardly lead us to conclude that a correlation might indicate a meaningful relationship between TOEFL scores and preparation effect on test performance.

Contrary to what learners expected, their performance on the reading section of the two valid and accredited international tests was not backed up by their previous repertoire of skills acquired in their secondary and tertiary studies. It would be concluded that that formal education of English in our schools and universities doesn't provide our English students with necessary and enough readiness for International and standard exams of English. So it is of high importance to have extra classes of skill teaching for English students as it is seen how effective and useful it was in for Non-English students. Last not the least is that this study was not a nation-wide, large scale and it is recommended that anyone willing to extend and make solid generalizations conduct it over a larger groups of learners with different levels of proficiency.

The findings of the study will provide evidence to the score validity that concerns university admission officers and professors, i.e. whether accurate inferences and decisions can be made from students' TOEFL scores to their actual English language proficiency. Finally, ETS research on test preparation has been more than an academic exercise. It has resulted in significant—even dramatic—modifications to several tests that ETS offers. These changes are perhaps the clearest example of the impact of ETS's research on test preparation. However, there have, arguably, been more subtle effects as well. Now, when new assessments are being developed, the potential coachability of proposed new test item types is likely to be a factor in decisions about the final composition of a test. Considerations about test preparation figure into the design of tests, well before these tests are ever administered to test takers.

REFERENCES

ACT. (2005). *What kind of test preparation is best?* Iowa City, IA: ACT web flyer retrieved March 14, 2007, from www.act.org.

Alderman, D. L., & Powers, D. E. (1980). The Effects of Special preparation on SAT- Verbal Scores. *American Educational Research Journal* 17 (2), 239-251.

Bachman, L. F., Davidson, F., Ryan, K., & Choi, I. C. (1995). An investigation of comparability of two tests of English as a foreign language. Cambridge: Cambridge University Press.

Bangert, R. L., Kulik, J. A., & Kulik, C. C. (1983). Effects of Coaching Program on Achievement Test Performance. *Review of Educational Research*, 53 (4), 571- 585.

Brown, J. D. H. (1998). An Investigation into Approaches to IELTS Preparation, with particular focus on the Academic Writing Component of the Test. In S. Wood (Ed.), *IELTS Research Reports* (Vol. 1, pp. 20-37).

Camilli, G., & Shepard, L. A. (1994). *Methods for identifying biased test items*. CA: Sage.

Celestine, C., & Ming, C. S. (1999). The Effect of Background Disciplines on IELTS Scores. In R. Tulloh (Ed.), *IELTS Research Reports* (Vol. 2, pp. 36-51). Canberra: IELTS Australia Pty Limited.

Chastain, K. (1988). *Developing Second-Language Skills: Theory and practice*. 3rd ed. Harcourt Brace Jovanovich, Publisher, Orlando, Florida: United State.

Cho, Y., & Bridgeman, B. (2012). Relationship of TOEFL iBT® scores to Academic performance: Some evidence from American universities. *Language Testing*, 29(3) 421-442. ltj.sagepub.com.

Elder, C., & K. O'Loughlin. (2003). Investigating the relationship between intensive English language study and band score gain on IELTS. In R. Tolloh (Ed.), *IELTS Research Report* (Vol. 4, pp. 207-254). Canberra: IELTS Australia Pty Limited.

Farnsworth, (2013) Effects of Targeted Test Preparation on Scores of Two Tests of Oral English as a Second Language. *TESOL Quarterly* Vol. 47.1, 148-186.

Geranpayeh, A. (1994). Are score comparisons across language proficiency test batteries justified?: an IELTS -TOEFL comparability study. *Edinburgh working papers in applied linguistics* 5, 50-65.

- Hamp-Lyons, L.(1991). Assessing second language writing in academic context. Norwood,N.G:Albex
- Hayes, B., & Read, J. (2004). IELTS preparation in New Zealand: preparing students for academic module. In L. Cheng, Y. Watanabe & A. Curtis (Eds.), Washback in language testing. Research contexts and methods (pp. 97-112). Mahwah: Lawrence Erlbaum Associates.
- Hayes, B., & Watt, L. (1998).An IELTS preparation course for Asian students; when practice doesn't make perfect. *EA Journal*, 16, 15-21.
- McLaughlin, Skaggs, and Patterson (2009). Preparation for and Performance on the GED Test. GED Testing Service®. Washington, DC: USA.
- McNamara, T., & Roever, C. (2006). Validity and the social dimension of language testing. *Language Learning*, 56 (Supplementary 2), 9-42.
- Messick, S. (1982). Issues of effectiveness and equity in the coaching controversy: Implications for educational testing and practice. *Educational Psychologist*, 17, 67-91.
- Miyasaka, J. R. (2000, April). A framework for evaluating the validity of test preparation practices . Presentation in the symposium, instructionally corrupt test preparation: Can it be detected or deterred? Paper presented at the Annual Meeting of the American Educational Research Association, April 2000, New Orleans, LA.
- Nguyen, T. (2007) Effects of Test Preparation on Test Performance the case of the IELTS and TOEFL iBT Listening Tests. *Melbourne Papers in Language Testing*. 1-23.
- Perlman, C. L. (2004). Practice tests and study guides: Do they help? Are they ethical? What is ethical test preparation practice? Chapter 27 in Wall, J. E., & Walz, G. R. (eds.). *Measuring Up: Assessment Issues for Teachers, Counselors, and Administrators*, 387-396. Greensboro, NC: CAPS Press.
- Powers, D. E. (1985). Effects of Coaching on GRE Aptitude Test Scores. *Journal of Educational Measurement*, 22 (2), 121 - 136.
- Powers, D. E. (1986). Relation of Test Items Characteristics to Test Preparation/ Test Practice Effects: a Quantitative Summary. *Psychological Bulletin*, 100 (1), 67 - 77.
- Powers, D., E. (2012). Understanding the Impact of Special Preparation for Admissions Tests. ETS R&D Scientific and Policy Contributions Series. ETS SPC-12-02 ETS Research Report No. RR-12-05.

- Rubenstein, J. (2004). Test preparation: What makes it effective? Chapter 28 in Wall, J. E., & Walz, G. R. (eds.). *Measuring Up: Assessment Issues for Teachers, Counselors, and Administrators*, 387-396. Greensboro, NC: CAPS Press.
- Sheshkelani, A., Ahari, H. & Aidinlou, N. (2012). An Investigation of Relationship between Test Preparation and Test Performance of Iranian EFL learners on Listening Skill in TOEFL. *International Journal of Basic & Applied Sciences IJBAS-IJENS*, 12, (2).
- Thiel, T. (1995). *An Analysis of the Evolution of the IELTS and an Investigation of Its Validity*. University of Tasmania, Hobart.