# A STUDY OF BACKWASH EFFECTS IN ENGLISH PROFICIENCY TESTING ON UNIVERSITY LOW-ACHIEVING STUDENTS

Mei-Sheng Chao

Chinese Culture University, Taiwan, ROC

Email: mschao@mail.sju.edu.tw

## Abstract

An understanding of backwash is increasingly important because more university students are required to pass standardized English proficiency tests before they graduate. It is well recognized that if a positive backwash effect can be fully adopted in the curriculum, then achieving good teaching quality and improving student learning is more likely. This study was particularly designed to see if a positive backwash effect can benefit low-achieving students, and help them regain their confidence in learning English. Eighty-three university students participated in this study and their learning achievements and strategies were examined. The course was an English remedial course which was used as an alternative to reach a threshold in English for graduation threshold, also a pre-test, post-test and questionnaire were used for quantitative analyses. In addition, individual interviews were applied to gather some qualitative data for more in-depth interpretation. Students were divided into three groups based on the pre-test scores: low, intermediate and advanced groups. The results show that, for low-achieving students, their post-test performance was superior to the other two groups (intermediate and advanced levels) in terms of listening and reading. No significant differences were found among the three groups in terms of learning strategies; however, individual interview data did show differences. Finally, the implications for designing such a course for low-achieving students in English are discussed and suggestions are proposed for future research.

***Keywords:*** Backwash, English Graduation Thresholds, Language Learning Strategy.

## 1. Introduction

Since 2004, the Ministry of Education in Taiwan encouraged local universities to set thresholds in English to be metin order to graduate, and most universities have cooperated in drawing up relevant plans to help students to fulfill this requirement. Universities have held workshops and short-period courses to strengthen students' capability for passing the proficiency tests, partly because they were used as the criteria for graduation thresholds. For

those who did not pass the tests, universities also provided a "make-up" formula, such as participating in extra "intensive English" or "remedial" courses, in order to get their graduation diplomas. For some college students, English proficiency tests are of great concern, not only because of constant failure, but also because of wasting a lot of money. Recently, denunciations of such a measure from students are becoming stronger. Quite a few students disagree with reaching a standard of English as a requirement for graduation.

Different from students, the attitudes of teachers toward English proficiency tests as graduation thresholds are more positive. Most teachers consider it necessary to set a graduation threshold for English, and their major concern focuses on students' competitiveness in the workplace after graduation. Although teachers recognize proper supplementary measures of English as graduation thresholds are necessary, they care more about the core issue, which is how to enhance low-achieving students' English to survive in the future. In order to help these students equip themselves with a good ability in English and to regain their confidence, it is important to understand their learning difficulties and apply relevant theories which can support course design and the effective use of English proficiency tests. The current study was particularly designed for this purpose. It is hoped that low-achieving students can benefit from a positive backwash effects in this 15-week course, where multiple materials were applied, including a textbook, several on-line proficiency practice tests, intra-group cooperation and support.

Pretest, posttest and Language Learning Strategy questionnaires were given at the beginning and end of the semester, and quantitative as well as qualitative data were collected for analysis. Specifically, this empirical study attempted to address the following questions:

1. Did low-achieving university students benefit from backwash effects, and show significant performance differences in studying English?
2. What are low-achieving students' difficulties in English learning and how to overcome them?

## 2. Literature Review

### Backwash Effect

In the process of learning, as a specific test becomes a standard for school entry or graduation standard, its backwash effect brings more attention to students, teachers, and educators. Ever since the Taiwanese government urged universities to set thresholds in English for graduation, English proficiency tests have played contradictory roles in English instruction. Some researchers indicate that the backwash effect of an English threshold for graduation is negative, while others propose that both teachers and students perceive a significant positive backwash effect. Proponents who argue that it has negative effects argue that, instead of

setting a threshold as a requirement for graduation, administrators should equip students with more power and knowledge to pass English Proficiency Tests. Improving teaching is considered to be one of the fundamental solutions to enhance students' English ability even if it is not achievable in the short term.

To help students, particularly low-achievers, reach the threshold, positive backwash can be used. If teachers can take advantage of it and embed the test contents and schemes into the curriculum, a positive backwash effect can be maximized and students' confidences will be regained.

The study carried out by Wall and Alderson in 1993 in Sri Lanka showed little positive backwash effect in an English intermediate proficiency test, but the tests did show a stronger impact on the content of courses rather than on teaching methods. Hamp-Lyons (1996) investigates the backwash effect in TOFEL test on teaching and finds that the backwash effect does exist in both contents and teaching methods. This study further points out that the teacher's personal factors, such as educational background, teaching experiences, etc., make big differences on the effect. Similar views are proposed by Watanabe, (1996) in a study examining the backwash effect on Japanese university students. Green, (2006) compares TOFEL and general writing classes through observation and interviews, and concludes that the backwash effect usually appears in the contents of English proficiency tests training, with much narrower contents. Meanwhile, he concurs with the view that the teacher's personal factors might have a greater influence on the backwash effect.

What most teachers really are concerned with is the necessity of rethinking the relationship between teaching, tests, and learning from the respect of the backwash effect. Some educators criticize the backwash effect for its focus on the potential impact of testing on learning; however, it is very useful for students to fully understand how these tests are structured to assess their learning outcomes. The primary concern of this study is to raise students' perceptions of the measuring schemes, and together with cooperative learning, to make teaching as a non-stop revising loop and to maximize the value of the backwash effect.
Strategy Inventory for Language Learning (SILL)

Since the1970s, various English as a Foreign Language (EFL) or English as Second Language (ESL) scholars have tried to find the characteristics and learning patterns of good language learners. Many local researchers have argued that through the appropriate management of learning strategies, students are more likely to reach their learning objectives.

Although some scholars define learning strategies as the learners' procedures for language acquisition and information access (Rigney, 1978), others (Weinstein & Mayer, 1986; Bialystok, 1978; Cohen, 1990,1998; MacIntyre, 1994; Oxford & Cohen 1992) consider learning strategies to be the learners' intentions which influence learners' behaviors, thinking and their coding procedures. To help students develop communicative skills through a more efficient, interesting and faster method, Oxford (1990) adopts these scholars' research results, and set up a classification system of English learning strategies. Broadly, these are divided into two categories: direct and indirect strategies. Under these two categories, he further classifies them into three sub-categories, including direct strategies: remembering techniques, cognition, and compensation strategies; and indirect strategies: metacognition, affection, and social interaction. While direct strategies relate to the target language learning techniques, such as memorizing, organizing, and smart guessing, etc., indirect strategies focus more on supplementary learning schemes, such as monitoring learning progress, reducing anxiety, and cooperating with others, etc. SILL (Strategy Inventory for Language Learning) was used in this study to explain how students perceive English learning and overcome their difficulties.

### 3. Methodology

**Participants**

Eighty-three university non-English major students participated in this study, with levels ranging from low to intermediate, based on proficiency test standards. All of them had been studying English for at least seven years, and had failed to meet the threshold for English set by the university to graduate.

Procedure

Multimedia instruction was applied in this study over the semester. Methods such as on-line practice tests, video watching, Powerpoint presentations, oral lectures, and the like, were all employed as teaching methods. Peer interaction and cooperative learning were especially emphasized.

Students were asked to log onto the web-enhanced course system, as shown in picture 1. They were also required to take listening quizzes (picture 2) every two weeks, and more listening practice sessions were provided (picture 3). One more requirement was to ask students to take the simulated English proficiency online tests every three weeks (picture 4). Students were divided into three groups based on their pretest scores and they were assigned to low, intermediate and high groups respectively. To take the advantage of cooperation and peer-learning, the advanced group of students was encouraged to help low group of students. The assistance provided by the advanced students included feedback on online tests and time management for study. At the end of the semester, all students took a posttest and filled out SILL questionnaire (Strategy Inventory for Language Learning). To obtain more in-depth data of students' learning strategy, face-to-face interviews were held with 12 students for

further interpretation. All these data were collected and SPSS was used further analyses.
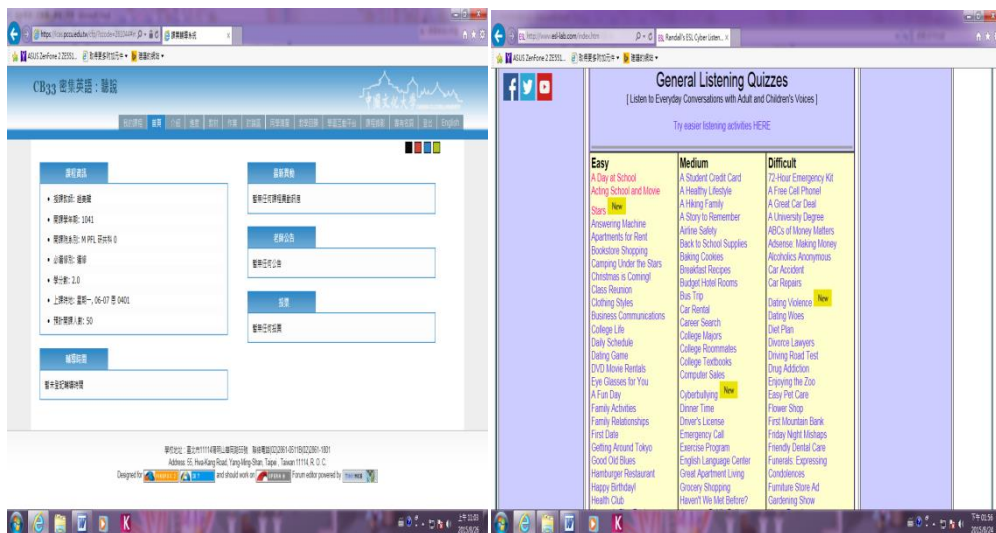

Picture 1 Web-enhanced learning system


Picture 2Online Listening Quizzes
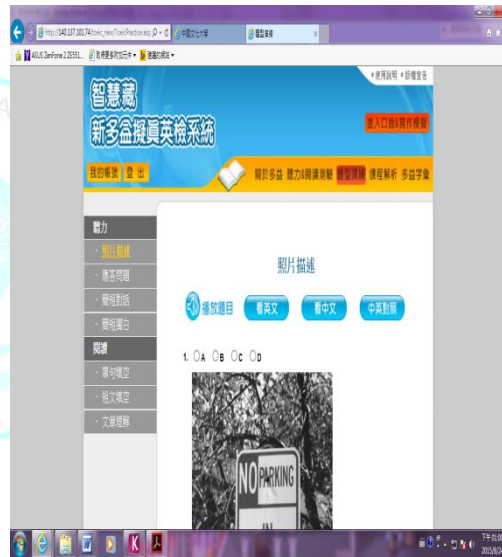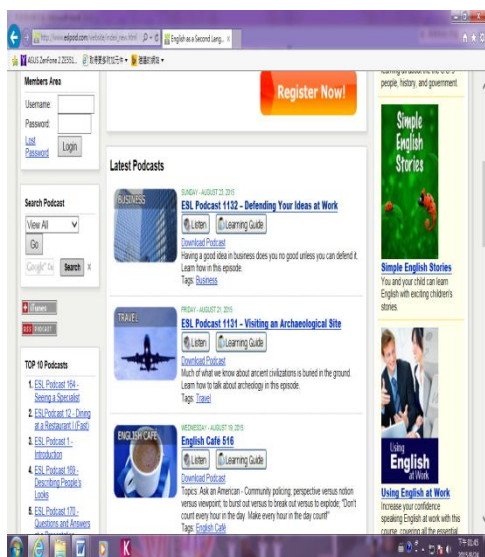

Picture 3Online Listening Practices


Picture 4Simulated proficiency tests

Instruments

To ensure the validity and reliability of the test, we chose English proficiency tests similar to TOEIC as the pretest and posttest; meanwhile the Strategy Inventory for English Learning (SILL) was used to understand students' English learning strategies. Both tests and questionnaires were well established and had a reputation of stable reliability and validity.

## 4. **Results**

Eighty-three university students participated in this study, and the pretest was given in groups. Three groups were divided using a normal curve, where the low and high groups consisted of 21 students respectively, and the other 41 were in the intermediate group. Analysis of covariance (ANCOVA) was applied to ensure a more precise statistical result, and Scheffe post hoc was used for further examination. Table 1 shows the original mean and the

adjusted mean of the listening part. It is found that for the low group, their improvement discrepancy is much larger than for the other two groups. In Table 2, the regression coefficients result did reach statistical significance level, where $F_{(2, 80)}=7.20$, $p<0.5$, which explains that the differences between groups were significant.

Table 1: Adjusted mean of listening of three groups

| Condition | Pretest Mean | Posttest Mean | Adjusted mean |
|---|---|---|---|
| low group | 25.67 | 32.95 | 35.63 |
| intermediate group | 34.12 | 39.29 | 39.245 |
| advanced group | 41.90 | 45.86 | 43.29 |

Note: condition refers to grouping

Table 2: Results of regression coefficients on listening scores

| variables | F | df1 | df2 | Sig. |
|---|---|---|---|---|
| listening score | 7.20 | 2 | 80 | .001* |

*p<.05

Table 3 presents the results of Scheffe post-hoc multiple comparisons. Significant differences were found among three groups ($F=6.45$, $p<0.5$), with the low group outperforming the other two groups in the listening part. It is very obvious that the improvement of the low group in listening comprehension is statistically significant.

Table 3:  Scheffe post-hoc multiple comparisons of listening scores

| Sources of variances | SS | df | MS | F | p | Post Hoc |
|---|---|---|---|---|---|---|
| pretest | 165.58 | 1 | 165.68 | 9.41 | .003 | |
| condition | 226.95 | 2 | 113.48 | 6.45 | .003 | 3>2*、3>1* |
| error | 1309.43 | 79 | 17.60 | | | |
| Corrected total | 3304.87 | 82 | | | | |

*p<.05

Note: condition refers to grouping; 3 refers to the low group

Regarding the reading comprehension part, there were no statistical differences found for the detailed mean scores in table 4, and table 5 shows the regression coefficients on reading scores, explaining why the F value did not reach statistical significance.

Table 4: Adjusted mean of reading for the three groups

| Condition | Pretest Mean | Posttest Mean | Adjusted Mean |
|---|---|---|---|
| Low | 21.71 | 24.52 | 30.16 |

| | | | |
|---|---|---|---|
| Intermediate | 31.46 | 32.49 | 32.15 |
| Advanced | 39.05 | 39.86 | 34.88 |

Table 5:Results of regression coefficients on reading scores

| Name of variables | F | df1 | df2 | Sig. |
|---|---|---|---|---|
| reading score | 0.23 | 2 | 80 | .797 |

*p<.05

Table6: Comparison between the Pretest and Posttest total scores

| Variance | N | Mean | Std. Deviation | t |
|---|---|---|---|---|
| Pre-test score | 83 | 64.87 | 13.31 | -8.98* |
| Pro-test score | 83 | 71.69 | 12.74 | |

*p<.05

A t-test was used to examine the statistical differences between pretest and posttest. According to the statistical analyses, the mean score of posttest was apparently higher than that of the pretest. The t value was -8.98, showing that there was a statistically significant difference between these two tests. Based on the details in Table 6, it can be inferred that students did much better in the posttest.

Table 7: Mean scores on each strategy of students

| Variances of strategies | N | Mean | Std. Deviation | t |
|---|---|---|---|---|
| retention | 83 | 3.54 | .89 | * |
| cognition | 83 | 3.54 | .88 | * |
| compensation | 83 | 3.66 | .87 | * |
| meta-cognition | 83 | 3.66 | .90 | * |
| Affection | 83 | 3.52 | .91 | * |
| interaction | 83 | 3.42 | .99 | * |

*p<.05

Table 7 shows the mean score and standard deviation of each sub-category in language learning strategy. The scale used in this questionnaire is a Likert scale, from 5 (strongly agree) to 1 (strongly disagree). Students' responses tended to be neutral to some extent. The most frequently used were metacognition and compensation strategies, and the least one used was an interaction strategy. In order to find more in-depth reasons for language learning

strategies, some verbal data were gathered through interviews, and these will be discussed in the following section.

## 5. Discussion

In general, this study showed that the backwash effect, when combined with cooperative learning with peers, did facilitate low-achieving students' learning achievement. As described previously, a positive backwash effect can raise student awareness and focuses on those knowledge and skills relevant to their assessments. In Taiwan, most universities require students to meet a threshold in English to graduate before getting a degree, and the most common and acceptable assessment schemes are English proficiency tests. It is well recognized that the ultimate goal of education is to equip students with the proper knowledge and skills relevant to their future, so if the test contents can be viewed as constructive rather than isolated on its own, then the positive backwash effect can be maximized. This study combined several techniques, including the backwash effect, cooperative learning, learning strategy scaffolding, and so on, to improve low achievers' English ability and build up their confidence. To answer the first research question, low-achieving students did benefit from a positive backwash effect, and their listening ability significantly improved. Although the improvement could be attributed to cooperative learning as well, nevertheless, constant practice and the familiarity of test structures did also help a lot. Regarding the reading part, although not statistically significant, low-achieving students still noticeably improve compared with the other groups.

It appears from the results of the study that low achievers did much better in listening than reading. The reasons are perhaps twofold. One is that students did more listening practice than reading, and the other concerns an increase in ability. Apparently, it takes much longer for students to build up their reading ability, unlike listening, which can be improved within a short period of time.

The second research question focuses on the language learning strategies which students use to learn English. The quantitative data shows no significant difference among different sub-categories; however, through individual interviews, some information was worth considering and provides valuable input in teaching.

For low-achieving students, they simply memorized new words as much as they could, and occasionally, they would use synonyms or a root word to help, but this was not effective. Their vocabulary accumulated slowly compared with intermediate and advanced level students, who were able to adopt more logical ways to memorize new words, such as taking the word apart, connecting them with different situations, etc.

Concerning the cognitive processes in learning English, all three groups would take advantage of cyber learning networks to correct their pronunciation and check word usage. Intermediate and low-achieving students would watch video clips with Chinese subtitles to

better understand the details of the contents. As for compensation for missing knowledge, all three groups frequently used "gesture" as a supplement to complete what they wanted to express. And "guess" became another favorite for students when interacting with foreigners. This situation occurred more frequently for low-achieving students, and once they figured out the meaning, they could apply easier words to fill in the blanks.

All three groups did not seriously monitor and evaluate their progress while learning English. Low-achieving students in particular were not autonomous learners. However, the advanced and intermediate groups did pay attention when others were speaking English, and tried to compare their own performances. For low-achieving students, job-related English requirements are the strongest incentive for them to monitor English learning.

To manage their emotion and anxiety, most advanced level students try to relax themselves, while the intermediate groups were stressed and afraid of making grammatical mistakes when learning English. On the other hand, low-achieving students were not afraid of making errors, even though they did make quite a few mistakes. When interacting with others in English, advanced students were keen to chat with foreigners, while the other two groups were more timid in doing so.

## Conclusion

Understanding how backwash works is increasingly important because more students are requested to take standardized English proficiency tests as one of the requirements for graduation. Over the last few decades, backwash studies have become even more critical in the field of language teaching because a lot of research points to the critical role of these high-stakes tests for students, teachers and society. These test results affect learners with different language proficiency levels, especially low-achieving students, partly because tests provide a clear focus to diagnose students' problems.

This study employed a cooperative learning strategy, where higher achieving students would offer assistance to lower level students. To ensure low-achieving students acquire enough basic proficiency test knowledge, the instruction of both teachers and peers were equally important in this study. Feedback from the online tests had very critical impact on their improvement. It is very obvious that when combined with other appropriate methods, the backwash effect becomes positive. Although learning should be continuous and not so narrow-focused, only by solving the short-term problems first can we encourage students to move on and broaden their views and knowledge. Research shows that for high-achieving students, they tend to use strategies more often than low-achieving ones (Green, 1991; Mullin, 1992; Park, 1994), which is the main reason for the author using cooperative learning in this study. Peer interaction leads to much greater influence on student learning achievements than the instructor. It is emphasized that learning English is a lifelong process, and many adults may not possess the learning strategies to enhance ability. The author strongly

suggests that cooperative learning should be embedded in courses, and low-achieving students need more attention to help them prepare for the competitive employment market.

## Acknowledgement

# References

i. Alderson, C. & Hamp-Lyons, L., 1996. TOEFL preparation courses: a study of washback. *Language Testing,* 13(3), pp. 280-297.

ii. Biggs, J., 1995. Assumptions underlying new approaches to assessment. *Curriculum Forum,* 4(2), pp. 1-22.

iii. Green, A., 2006. Watching for washback: observing the influence of the international English language testing system academic writing test in the classroom. *Language Assessment Quarterly,* 3(4), pp. 333-368.

iv. Green, J., 1991. *Language learning strategies of Puerto Rican University students. Paper presented at the annual meeting of Puerto Rico at Mayaguez,* Mayaguez: Puerto Rico University.

v. Hsu, L. & Wang, P., 2006. A study on English graduation threshold and English Learning Motivation. *Journal of National Kaohsiung University of Applied Sciences,* Volume 35, pp. 51-68.

vi. Mulllins, P., 1992. *Successful English language learning strategies of students enrolled in the faculty of arts,Chulalongkorn University, Bangkok, Thailand. Unpublished doctoral dissertation,* San Diego: U.S. International University.

vii. Oxford, R., 1990. *Language learning strategies: What every teacher should know.* New York: Newbury House Publisher.

viii. Park, G., 1994. *Language learning strategies: Why do adults need them?.* Austin: University of Texas.

ix. Sung, Y., Chang, K., Lee, Y. & Yu, W., 2008. Effects of a mobile electronic guidebook on visitors' attention and visiting behaviors. *Educational Technology and Society,* 11(2), pp. 67-80.

x. Sy, B., 1995. *Gender differences, perceptions on foreign language learning and language learning strategies.* Taipei, Taiwan, 12th National Conference on TESOL.

xi. Tsai, Y. & Tsou, C., 2009. A standardized English language proficiency test as the graduation benchmark: Student perspectives on its application in higher education.. *Assessment in Educaion:Principles, Policy & Practice,* 16(3), pp. 319-330.

xii. Watanabe, Y., 1996. Does grammar translation come from the entrance examination? Preliminary findings from classroom-based research. *Language Testing,* 13(3), pp. 318-326.

xiii. Yang, N., 1993. *Understanding Chinese students' language beliefsand learning strategies use.* Atlanta, GA, Annual meeting of International Teachers of English to speakers of other languages.

xiv. Yang, N., 1994. *An investigation of Taiwanese college students' useof English Learning Strategies,* Taiwan: National TaiwanUniversity.