# 外国語としての英語教育カリキュラムにおけるCALLの有効性についての発展的研究

デイヴィッド・レヴィン

インターネット検索活動の有効性を示す筆者の先行研究成果に基づき，本研究は外国語としての英語教育においてこの種の学習活動を支持するさらなる証拠を示すものである。本研究のため，英語CALL授業に登録した3クラス，89名の学生対象にウエブ中心の授業の最初と最終15週目に英語学力評価測定を行った。15週の学習により，概して，学生は平均4.8ポイント向上した。また，各クラスとも同様な平均値を示す向上が見られた。このような数値は統計的に優位性を示すものであり，その教育効果は，インターネットを利用した教育の有用性を示すものとなったといえる。

# The Effectiveness of CALL in an EFL Curriculum: A Follow-up Study

David Levin

## Introduction

In 2001, a colleague and I conducted a study to determine the effect, if any, the inclusion of CALL (Computer Assisted Language Learning) would have on an EFL curriculum (Levin & Redfield, 2003). The results were generally supportive of the particular instruction employed. Since that time, more than three years later (the study was published a year and a half later), great strides have been made in Japanese universities (and in many high schools) to upgrade computer facilities used for class and independent instruction. With servers upgraded and high Internet connections established throughout most parts of Japan, use of the Internet has become an increasingly popular and efficient tool for English instruction and research. In addition, instructors, both non-Japanese and Japanese who have embraced this important tool for their classes, are in demand at the various institutions across Japan. In fact, these days it is hard to find an employment advertisement for university English instructors that doesn't mention a desire for computer and Internet ability. With the ubiquitous use of the Internet firmly established in Japanese universities, tailoring its use to provide effective instruction in the language classroom continues to be a goal many instructors are striving towards. It is believed that the curriculum used in the above study is just such an effective use of the Internet. Based on the encouraging results of the previous research, this present study aims to replicate and build upon the established evidence that supports the use of

the Internet in the EFL classroom.

## Background

### *Web-based Language Learning*

In the previously mentioned study examining the inclusion of CALL in an EFL curriculum, Levin and Redfield tested 68 business and economics learners from a private university prior to and immediately after a course in which the students were taught using a curriculum that Lassche (2000) describes as "Web-based Language Learning." Results of the study showed that after a semester of instruction (12 weeks of total instruction), students gained an average of over three points on a 57 questions reading measures test. Even though the students in question were taking the CALL course as part of a broader English curriculum, positive effects of Web-based instruction seemed to be supported by the results.

This present study replicates the above study by, once again, using the same Web-based curriculum component as the medium of instruction, and also by testing the subjects prior to and after instruction, using the same reading measures test. However, for this current study, the period of instruction lasted a bit longer, approximately 15 weeks. The subjects for this study included 89 engineering and science students in one first-year and two second-year undergraduate English courses at Toyohashi University of Technology; the students were assigned to these classes based on English ability. As in the previous study, this study attempted to answer the question: Would there be meaningful measurable progress as measured by a standardized, multiple-choice test of general proficiency? In addition, an answer to another question was sought: Would there be meaningful measurable progress in each of the three classes evaluated separately?

### *Rationale*

Although research into the use of the Internet in language learning has been encouraging (e.g., Fox, 1997; Kimball, 1998; Hoshi, 2002), the results from Levin and Redfield (2003) study were only the first set of data to support the particular method of Web-based instruction used in that analysis. More

studies are needed to back up these preliminary results. In addition, students from this previous study likely benefited from the additional English courses taken in addition to the CALL course. Therefore, the effect of the Web-based instruction alone is difficult to assess. With the present study, the students were only taking one additional English course; therefore, the attribution of Web-based instruction can be given more weight. How would this fact be reflected in the data? Moreover, since the period of instruction was slightly longer, would the data reflect this change by showing greater gains in English ability?

Another facet not looked at in the previous study was, evaluated separately, how would individual classes perform. Considering the different years and different English ability of the three classes in this study, one could expect class dynamics to differ somewhat. Would any of these classes not respond well to this type of CALL instruction?

## Method

### *Participants*

89 students from Toyohashi University of Technology (TUT), enrolled in three separate CALL English classes, participated in this study. The English courses at TUT are streamed according to the results of a placement test. First and second-year students are placed into one of three class levels that are given a letter designate A, B or C ('A' being the highest level). Out of the 89 participants, 34 participants were enrolled in the second-year English 'B' course; 24 participants were enrolled in the second-year English 'C' course; and 31 participants were enrolled in the first-year English 'B' course. The students all were majoring in a variety of science and technology disciplines. Finally, only the results of those participants taking both the pretest and the posttest were included in this analysis.

### *Instruction*

The classes were content-based courses built around the Internet using Web-searching exercises. Learners were assigned specific questions to answer via self-initiated Web searches in each lesson. Neither grammatical

structure nor vocabulary was taught as such, although, of course, the instructor was present at all times to answer questions related to vocabulary, meaning and procedures. In essence, the students worked individually, with no formal, whole class lectures. The students worked at their own pace, yet within a limited time frame of generally 65 minutes. For the purposes of this paper, instruction is not a variable since there is no control group. We are simply trying to see, statistically, if two semesters of EFL, using CALL (Web-based) as a facet of the curriculum, leads to meaningful progress in EFL.

### *Instrument*

The Matsushita Pilot Placement test (MAT) parts I-III, a multiple-choice exam containing structure, vocabulary, reading, and cloze sections was chosen as the measuring instrument for this study. The MAT is an aptitude measure written specifically for post-secondary Japanese learners of English.

### *Administration*

The regular classroom English teacher administered the MAT during the second week of the second trimester in September 2003, and then again during the last week of the third trimester in February 2004. Each trimester at TUT lasts 9 or 10 weeks. Forty minutes were allowed for completion of the test.

### *Statistical Analysis*

InStat 3 for the Macintosh was used to derive descriptive statistics and to run the Kolmogorov and Smirnov Normality Test necessary to gauge the Gaussian distribution. The alpha for statistical significance was set at .05.

### *Research Questions*

1. Combining all participants in one group, will there be measurable progress in English acquisition over the course of two trimesters, as measured by the MAT?

2. Broken down separately, will each of the three classes also show measurable progress in English acquisition over the course of two trimesters, as measured by the MAT?

## Results

### *Overall proficiency*

For the pretest, the combined classes (n=89) had a mean score on the MAT of 33.730 (out of 60) with a standard deviation of 10.463 and a standard error of 1.109. The minimum and maximum scores were 7 and 56 respectively; the median score was 35. Finally, the lower and upper 95% confidence intervals (CI) were 31.523 and 35.938 respectively. As for the posttest, the mean score on the MAT was 38.539 with a standard deviation of 7.390 and a standard error of .7834. The minimum and maximum scores were 20 and 56 respectively; the median score was 39. Finally, the lower and upper 95% CI were 36.980 and 40.099 respectively (see Table 1 for descriptive statistics).

**Table 1**

***MAT Descriptive Statistics for Overall Pretest and Posttest***

| Parameter: | Pretest | Posttest |
|---|---|---|
| Mean: | 33.730 | 38.539 |
| # of points (N): | 89 | 89 |
| Std. deviation: | 10.463 | 7.390 |
| Std. error | 1.109 | .7834 |
| Minimum: | 7 | 20 |
| Maximum: | 56 | 56 |
| Median: | 35 | 39 |
| Lower 95% CI: | 31.523 | 36.980 |
| Upper 95% CI: | 35.938 | 40.099 |

For the combined classes, the data passed the Kolmogorov and Smirnov (KS) test for determining Gaussian distribution; the results yielded a KS distance of .07. In addition, because the data came from independent pairs, a statistically powerful paired t-test was chosen to further analyze the data. The test revealed an extremely significant two-tailed P value less than .0001 (t=7.282 with 88 degrees of freedom). Moreover, with the correlation coefficient (r)=.8101, the pairing for these data appears to be effective. Therefore, the observed difference between the two means from the pretest and the posttest is statistically extremely significant; the participants had higher posttest scores on the MAT (see Figure 1).

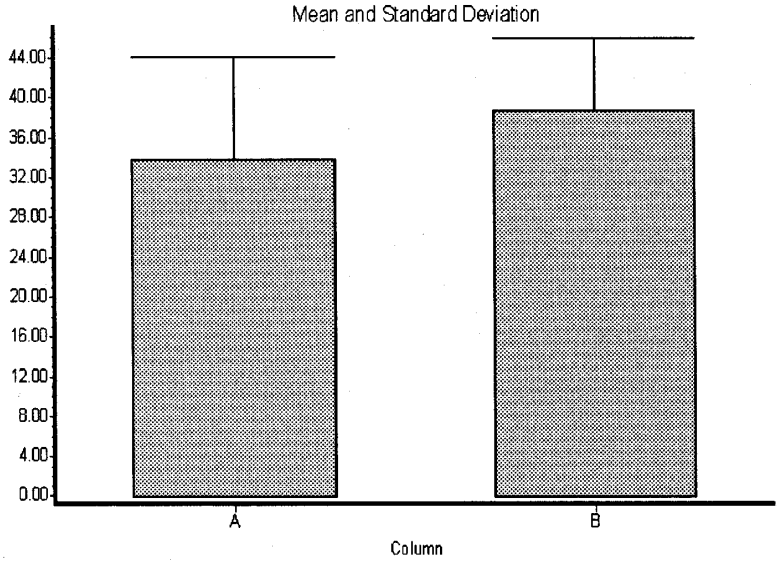


*Figure 1.* Overall mean and standard deviation for pretest and posttest (A=pretest, B=posttest).

### *Second-year intermediate class proficiency*

For the pretest, the second-year intermediate class (n=34) had a mean score on the MAT of 32.706 (out of 60) with a standard deviation of 9.631 and a standard error of 1.652. The minimum and maximum scores were 7 and 53 respectively; the median score was 34.5. Finally, the lower and upper 95% confidence intervals (CI) were 29.343 and 36.068 respectively. As for the posttest, the mean score on the MAT was 37.794 with a standard deviation of 7.405 and a standard error of 1.270. The minimum and maximum scores were 20 and 56 respectively; the median score was 38.5. Finally, the lower and upper 95% CI were 35.209 and 40.379 respectively (see Table 2 for descriptive statistics).

**Table 2**

***MAT Descriptive Statistics for Second-Year Intermediate Pretest and Posttest***

| Parameter: | Pretest | Posttest |
|---|---|---|
| Mean: | 32.706 | 37.794 |
| # of points (N): | 34 | 34 |
| Std. deviation: | 9.631 | 7.405 |
| Std. error | 1.652 | 1.270 |
| Minimum: | 7 | 20 |
| Maximum: | 53 | 56 |
| Median: | 34.5 | 38.5 |
| Lower 95% CI: | 29.343 | 35.209 |
| Upper 95% CI: | 36.068 | 40.379 |

For the second-year intermediate class, the data passed the Kolmogorov and Smirnov (KS) test for determining Gaussian distribution; the results yielded a KS distance of .09. In addition, because the data came from independent pairs, a statistically powerful paired t-test was chosen to further analyze the data. The test revealed an extremely significant two-tailed P value less than .0001 (t=5.756 with 33 degrees of freedom). Moreover, with the correlation coefficient (r)=.8485, the pairing for these data appears to be effective. Therefore, the observed difference between the two means from the pretest and the posttest is statistically extremely significant; the participants had higher posttest scores on the MAT (see Figure 2).

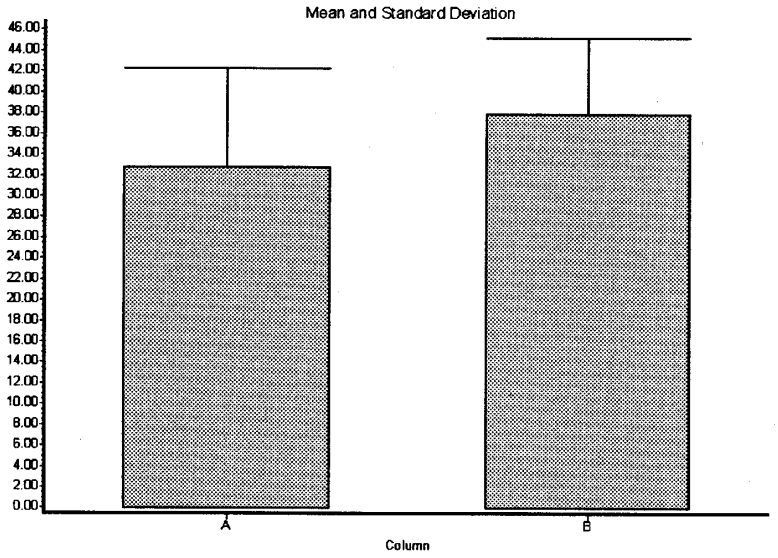


*Figure 2.* Second-year intermediate class mean and standard deviation for pretest and posttest (A=pretest, B=posttest).

### *Second-year lower-intermediate class proficiency*

For the pretest, the second-year lower-intermediate class (n=24) had a mean score on the MAT of 29.542 (out of 60) with a standard deviation of 10.616 and a standard error of 2.167. The minimum and maximum scores were 12 and 52 respectively; the median score was 30.5. Finally, the lower and upper 95% confidence intervals (CI) were 25.058 and 34.025 respectively. As for the posttest, the mean score on the MAT was 34.917 with a standard deviation of 5.926 and a standard error of 1.210. The minimum and maximum scores were 23 and 47 respectively; the median score was 34. Finally, the lower and upper 95% CI were 32.414 and 37.420

respectively (see Table 3 for descriptive statistics).

**Table 3**

***MAT Descriptive Statistics for Second-year Lower-Intermediate Pretest and Posttest***

| Parameter: | Pretest | Posttest |
|---|---|---|
| Mean: | 29.542 | 34.917 |
| # of points (N): | 24 | 24 |
| Std. deviation: | 10.616 | 5.926 |
| Std. error | 2.167 | 1.210 |
| Minimum: | 12 | 23 |
| Maximum: | 52 | 47 |
| Median: | 30.5 | 34 |
| Lower 95% CI: | 25.058 | 32.414 |
| Upper 95% CI: | 34.025 | 37.420 |

For the second-year lower-intermediate class, since the sample size was lower than 30 (N=24) a nonparametric, a Wilcoxon matched-pairs signed-ranks test was used to analyze results. The test revealed a very significant two-tailed P value of .0033. Moreover, with a nonparametric Spearman correlation coefficient (r)=.7515, the pairing for these data appears to be effective. Therefore, the observed difference between the two means from the pretest and the posttest is statistically very significant; the participants had higher posttest scores on the MAT (see Figure 3).

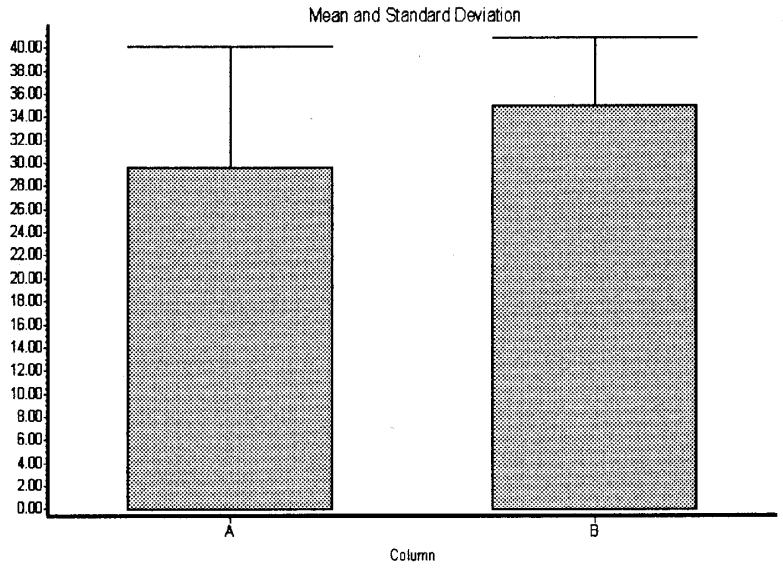


*Figure 3.* Second-year lower-intermediate class mean and standard deviation for pretest and posttest (A=pretest, B=posttest).

### *First-year intermediate class proficiency*

For the pretest, the first-year intermediate class (n=31) had a mean score on the MAT of 38.097 (out of 60) with a standard deviation of 9.887 and a standard error of 1.776. The minimum and maximum scores were 22 and 56 respectively; the median score was 37. Finally, the lower and upper 95% confidence intervals (CI) were 34.471 and 41.723 respectively. As for the posttest, the mean score on the MAT was 42.161 with a standard deviation of 6.948 and a standard error of 1.248. The minimum and maximum scores were 29 and 55 respectively; the median score was 42. Finally, the lower and upper 95% CI were 39.613 and 44.709 respectively (see Table 4 for descriptive statistics).

**Table 4**

***MAT Descriptive Statistics for First-Year Intermediate Pretest and Posttest***

| Parameter: | Pretest | Posttest |
|---|---|---|
| Mean: | 38.097 | 42.161 |
| # of points (N): | 31 | 31 |
| Std. deviation: | 9.887 | 6.948 |
| Std. error | 1.776 | 1.248 |
| Minimum: | 22 | 29 |
| Maximum: | 56 | 55 |
| Median: | 37 | 42 |
| Lower 95% CI: | 34.471 | 39.613 |
| Upper 95% CI: | 41.723 | 44.709 |

For the first-year intermediate class, the data passed the Kolmogorov and Smirnov (KS) test for determining Gaussian distribution; the results yielded a KS distance of .11. In addition, because the data came from independent pairs, a statistically powerful paired t-test was chosen to further analyze the data. The test revealed a very significant two-tailed P value of .0016 (t=3.473 with 30 degrees of freedom). Moreover, with the correlation coefficient (r)=.7538, the pairing for these data appears to be effective. Therefore, the observed difference between the two means from the pretest and the posttest is statistically extremely significant; the participants had higher posttest scores on the MAT (see Figure 4).

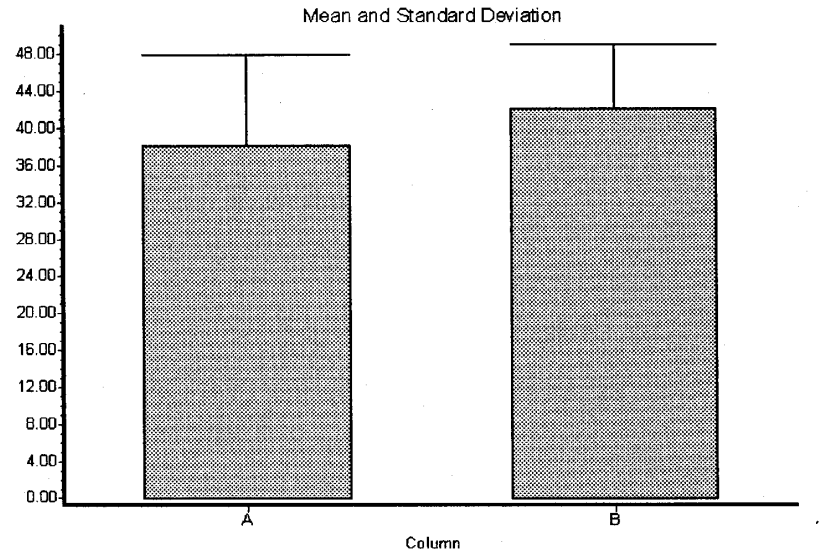


*Figure 4.* First-year intermediate class mean and standard deviation for pretest and posttest (A=pretest, B=posttest).

## Discussion

Supporting the results from Levin and Redfield's (2003) previous study, the results of this study also show support for the inclusion of CALL (Web-based) in an EFL curriculum. Overall, the students in this study improved an average of 4.8 points on the MAT over the course of two trimesters (approximately 15 weeks of instruction), progressing from a mean score of 33.730 to 38.539 on the 60-point English aptitude test. This difference was found to be statistically, extremely significant. Broken down into separate classes, we also see about the same amount of progress with the second-year intermediate students achieving a 5-point gain (extremely significant); the second-year lower intermediate students, 5.4 (very significant); and the first-year intermediate students gaining 4 points (very significant).

The overall gain of 4.8 points for the participants in this study is somewhat better than the 3 point gain made by the participants in the previous study. At first glance, this seems to be in line with the slightly longer period of instruction for this study's group. However, the subjects of this study were taking only one other English course per week as opposed to the three other taken by the subjects in the previous study. Moreover, the actual class period for the current study's students is 15 minutes shorter than that of the previous study. A quick calculation shows that during the period

of this study, the participants received a total of roughly 41 hours of English instruction, 19 of which was CALL. Compared with the previous study's participants of roughly 80 hours total English instruction, 18 CALL, the results shown in this study appear more dramatic. One explanation for this big jump in improvement might be explained by the overall improvement in computer literacy in Japan over the two-year period. Another might be the nature of the subjects themselves: science and technology students are more apt to be computer savvy than their business and economics counterparts, thus resulting in more effective use for language acquisition.

Another interesting aspect of the results is that each separate class, second-year intermediate and lower intermediate, and first-year intermediate, all had similar gains. It is important to note that the information sought out on the Internet during the day's activity in these three classes was different from class to class. It was the structure and general format of the activity that remained the same (the same format was also used in the previous study). In other words, the actual material that the students were required to seek out and learn was not as important as the format in which the information was sought out. It was the act of searching the Internet for information—reading blocks of English text in the process—that appears to have affected the results of this (and the previous) study. Stakhnevich (2002) lends support for this conclusion when she states, "... when teaching content through self-directed reading, Web medium can evoke better reading comprehension than traditional print medium."

Finally, as with the previous study, there are yet limitations to be addressed. While this study has greatly reduced the effect of other English courses on the results, the Web-based English instruction cannot be totally isolated has the primary reason for English proficiency improvement. In future studies on this topic, a more involved study that includes a control group would be desirable. However, the data presented in this study add to the growing body of evidence that supports the use of the Internet in EFL curriculum.

## Conclusion

In this study, 89 students enrolled in three separate English CALL (Computer Assisted Language Learning) courses were given an English ability

assessment measure before and after a 15-week period of Web-based instruction. As a whole, the students made an average gain of 4.8 points on the measure. These results were determined to be statistically, extremely significant. Similar statistically significant gains were also made when evaluating each of the classes separately. This study expanded upon and added to the body of evidence from a previous study (Levin & Redfield, 2003) showing that an EFL curriculum including courses using Web-based instruction can lead to student progress in English proficiency. Furthermore, in this study, the effects of this type of instruction were shown to be more greatly pronounced. It is clear from the growing body of research that use of the Internet and the Web is proving to be a boon for English language learning that fits in with the desire of educators to tap into Japan's advanced technological infrastructure.

## References

Fox, M. (1997). Beyond the technocentric: Developing and evaluating content-driven, Internet-based language acquisition courses. *Computer Assisted Language Learning, 10*(5), 442–453.

Hoshi, M. (2002). Practices, beliefs and perceptions of Japanese EFL self-access learners toward Internet-based language learning. *CALL-EJ Online, 4*(1). Retrieved June 11, 2002, from: http://www.clec.ritsumei.ac.jp/english/callejonline/7–1/hoshi.html.

Kimball, J. (1998). Task-based medical English: Elements of Internet-assisted language learning. *Computer Assisted Language Learning. 11*(4), 411–417.

Lassche, G. (2000). Web-based language learning in Korea: A pedagogical critique. *Korea TESOL Journal*, 3(1), 55–76.

Levin, D. & Redfield, M. (2003). A preliminary look into the effectiveness of the inclusion of CALL in an EFL curriculum. *Toyohashi University of Technology: The Lark Hill, 25*, pp. 67–75.

Stakhnevich, J. (2002). Reading on the Web: Implications for ESL professionals. *The Reading Matrix, 2*(2). Retrieved October 4, 2002, from: http://www.readingmatrix.com/articles/stakhnevich/.