

DOI: [10.20472/TE.2017.5.2.007](https://doi.org/10.20472/TE.2017.5.2.007)

HOW COMPUTER-BASED AND PAPER-BASED MODES OF LEARNING AFFECT STUDENT PERFORMANCE ON VOCABULARY TESTS IN THE TOEIC CLASSROOM

NICHOLAS STONE, YUSUKE HAYASAKA, CHIZUKO AIBA, REIKO YAMASHITA, JASON BARROWS

Abstract:

The authors of this paper examined how paper-based vs. computer-based modes of learning affected student performance on TOEIC vocabulary tests for known and unknown words. Japanese university students were tested on two different TOEIC vocabulary registers to see how the mode of learning affected student performance on vocabulary tests. The results show all students, including higher ability students, perform better on paper-based over computer-based tests and this performance is carried through all levels of ability until students become familiar with the vocabulary, and their cognitive load is decreased. This indicates that all students when challenged by new and unknown data would initially perform better with paper-based modes of learning. This has important implications for the teaching of new TOEIC vocabulary both at the institutional and classroom level.

Keywords:

TOEIC, computer-based modes of learning, paper-based modes of learning, learner preference, learning styles, test mode effect, vocabulary learning

JEL Classification: I21, I23

Authors:

NICHOLAS STONE, Tung Wah College, Hong Kong, Email: nick_stone_187@hotmail.com

YUSUKE HAYASAKA, International University of Health and Welfare, Japan, Email: hayasaka@iuhw.ac.jp

CHIZUKO AIBA, Tokyo Denki University, Japan, Email: aiba@mail.dendai.ac.jp

REIKO YAMASHITA, Seijo University, Japan, Email: englishyamashita@gmail.com

JASON BARROWS, Tokyo Denki University, Japan, Email: jasonb@mail.dendai.ac.jp

Citation:

NICHOLAS STONE, YUSUKE HAYASAKA, CHIZUKO AIBA, REIKO YAMASHITA, JASON BARROWS (2017). How Computer-based and Paper-based Modes of Learning Affect Student Performance on Vocabulary Tests in the TOEIC Classroom. *International Journal of Teaching and Education*, Vol. V(2), pp. 105-116., [10.20472/TE.2017.5.2.007](https://doi.org/10.20472/TE.2017.5.2.007)

Introduction

The authors teach compulsory English courses to first-year undergraduate students majoring in Information Technology. The university has been proactively adopting the latest technology, such as digital devices and multimedia in its classrooms, and for the past several years, English classes also have incorporated an e-learning environment with computers and internet access. All first-year undergraduate students have been required to purchase and bring their own computers to all of their English classes. Despite the institutional enthusiasm for e-learning, in which the authors shared, they began to become somewhat concerned when they began to perceive a pattern forming; when faced with new, unfamiliar and or difficult content, all level of learners appeared to perform better with traditional paper-based learning and testing. In other words, students appeared to be disadvantaged by computer-based learning.

At the beginning of a new semester, the authors introduced a new curriculum of work to the students, comprising the content and use of computers. After three weeks of experimental use with the computers, the authors of this paper took the opportunity to investigate how the students responded to the use of computer-based, and paper-based vocabulary learning in the classroom. The authors used a recognized vocabulary list provided by the Ministry of Education, of language level appropriate words for first-year university students to conduct weekly vocabulary tests. The research project, then, focused on the following questions: How do students respond to the use of computer-based vs. the paper-based learning models in English classes? Secondly, using these test scores, can an understanding of the relationship between language learning and the use of computer-based vs. paper-based learning modes be reached and how do these affect student performance? The ultimate goal of this study is to enhance language teaching by increasing our understanding on the appropriate adoption and use of new technologies in the classroom.

In preparation for this research, the authors read extensively into the development of e-learning (hereafter referred to as computer-based learning) and its strengths and weaknesses. What soon became apparent was that there was a dearth of the literature concerning a fundamental question; what benefits does computer-based learning bring over traditional modes of learning and where is the evidence that the former ultimately benefit student performance on learning. Therefore, the overreaching goal of this research changed to produce empirical evidence into this hitherto under-researched field of study, i.e. the benefits or not, of computer-based learning over traditional modes of learning. Furthermore, it was hoped that this research would sound a note of caution against the assumption that computer-based learning is unquestionably of benefit to all

learners and encourage institutions and educators to use this mode of learning as they would any other; a tool, that comes with its own pros and cons, to be used accordingly.

Literature Review

The use of new technologies in language education is not new. Since the 1960s, CALL (Computer Assisted Language Learning) has played a significant role in language teaching, being theoretically supported by the progress of second-language acquisition theory (Gruba, 2004; Levy & Stockwell, 2006). More recently, since the mid-1990s, researchers have been pursuing the possibility of MALL (mobile-assisted language learning) by exploring the use of mobile devices such as tablets, MP3 players, handheld gaming devices, and mobile phones in language education (Kukulka-Hulme, & Shield, 2008). However, compared to other learning aids, there is a paucity of empirical research on the use of technology in language learning (Low, 2014, p. 165).

The last ten years or so has seen a marked increase in the institutional adoption and use of technology in the classroom as the cost and size decreased and became markedly more available. Various benefits have been associated with this adoption such as tailoring to learner differences, increasing learner motivation, ease of administration of computer-based tests, etc. However, the recent enthusiasm for the adoption of computer-based learning has not in fact been based on a similar abundance of research that has proven a fundamental assumption that computer-based learning provides for similar or better student performance over traditional methods (Toth 2008). Yet few authors have sounded a note of caution over the adoption of computer-based learning, with the noticeable exception of Emerson and MacKay (2011, pp. 734) who have written that educators

...need to consider how best to develop teaching and learning strategies that incorporate the strengths and opportunities presented by online learning with the strengths of more traditional modes of learning. And central to this exploration must be the enhancement of our students' learning outcomes.

In second language acquisition theory (SLA), various aspects have been suggested as affecting student performance such as sex, age, ability, preferred learning style and so on. A further, under appreciated aspect, is called cognitive load theory. In the late 1980s, cognitive scientist John Sweller revealed in his study of problem-solving that learning is inhibited if learners have too high a cognitive load on their working memory (Sweller, 1988). Cognitive load on working memory is divided into three types: intrinsic,

extraneous, and germane (Paas, Renkl, & Sweller, 2004). Intrinsic load is a fundamental element and cannot be lowered, while extraneous load can be decreased by instructional designs. When people describe a triangle, for example, explaining a shape only by words is more difficult than showing a picture. In this case, the picture decreases extraneous load. After changing the level of extraneous load, the cognitive load for the task does not exceed the limit of working memory. Germane load is said to encourage and automate schematization, which occurs when knowledge of short-term memory is converted into long-term memory.

Previous research has positively linked learner's preferred learning style and performance. Wallace and Clariana (2000) in their study found that learners who were competitive, unfamiliar with the content and computers did better with the paper-based test. The opposite was true with high-attaining students who did better with computer-based tests (Watson, 2001; Clariana & Wallace, 2002). Nevertheless, if we take competitiveness as a trait of extroversion, then it could be predicted that extroverts will perform better on paper-based tests and the opposite with computer-based tests and introverts.

Various studies have been conducted examining the difference in student performance and computer-based vs. paper-based learning and testing. However, the results have been mixed and inconclusive with some studies finding in favor of better student performance for one or the other mode, while others have found no difference between the two. One possible cause for such inconclusive results could be due to the different use of terms and often overlap meaning. It is difficult to write absolute definitions of what types of learning styles exist, making comparisons between studies problematic (Hall, 2011). Additionally, it has been argued that there may be an issue of equivalence. Despite the widespread and enthusiastic adopting of computer-based learning, comparative little in-depth research has been conducted as to how this mode of testing may affect student performance as opposed to traditional modes, e.g. the effects on the cognitive workload, stress or preference of the test taker (Noyes, Garland & Robbins, 2004). Leeson (2006) writes that the ability to process information, computer familiarity and anxiety could be uniquely related to computer-based test mode effect. Learners when tested on unfamiliar content, working under increased stress and cognitive load levels, may be disadvantaged when using computer-based tests as compared to traditional and more familiar paper-based assessments, unfairly affecting their performance. With regards to preference, prior positive experience with previous computer-based tests results in more positive attitudes to them (Fluck, Pullen & Harper, 2009).

In summary, rather than one test mode providing for better student performance for all ability of learners, previous research suggests that preferred learning style, test mode effect, test mode preference, and familiarity with the content play an important role in determining student performance. Despite this, it has been assumed and become a widely-adopted ethos that computer-based learning has an advantage over traditional paper-based modes of learning and assessment for the majority of students. However, the authors questioned this based on their experience and observation of their own student's performance and decided to conduct a study investigating this situation. This study collected two data sets, computer-based and paper-based learning, and compared the difference in performance between the two sets.

Method

In their classes, the authors administered two forms of instruction, one computer-based and one paper-based, on alternating weeks to their students. At the end of each week, the students were administered twelve weekly vocabulary tests, which consisted of ten questions, each on an individual vocabulary item. The students were given a list of twenty-five possible vocabulary items and forty-five minutes to study them immediately prior to being tested upon them. This was where the effect of either mode of learning was expected to be realized.

Two separate vocabulary lists were used which represented two registers; general and medical English words. The first represented words that the students are expected to be familiar with as a result of their general English education from junior high and high school. Medical was chosen for its expected unfamiliarity amongst the students; such an unfamiliar register was chosen to test the hypothesis as to whether or not it was the mode of learning or familiarity with the content that determined performance. The vocabulary was selected from a list of common use TOEIC vocabulary. The study consisted of a group of thirty-seven exclusively first-year Japanese engineering university students; five of whom were female, and thirty-two were male.

Results and Discussion

The analysis showed that despite the use of the TOEIC's approved vocabulary list, students performed 39% better on general vocabulary items as compared to medical. These results are interesting due to the fact that while all of these vocabulary words appear on the Ministry of Educations list, the medical scores were significantly lower. One possibility for this result is that a non-medical student would not be interested in this registry of words, and as a result, their preference would not be actualized. These two sets of results are not surprising because they represent, first of all, vocabulary items that learners had been exposed to since their compulsorily high-school education.

In the latter case of medical words, it can be supposed that this registry is relatively new to the students and thus even given the forty-five minutes in which to learn the vocabulary, their germane load remained quite high, thus increasing their cognitive load, negatively affecting performance.

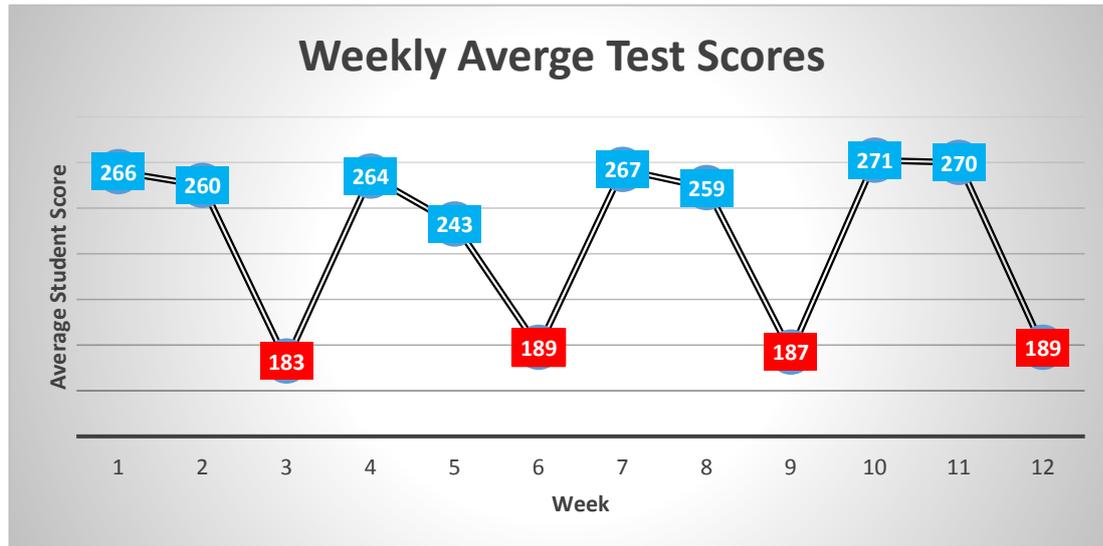


Figure 1: Chart is showing the average test scores for all students. ● Medical Register
● General Register. Odd weeks = paper-based learning. Even weeks = computer-based learning.

Looking at chart 1 below (pp. 7), it should be noticed that all the students' scores were lower for weeks three, six, nine and twelve, the medical weeks and that no students were able to get a perfect score. This suggests that rather than it being an issue of student ability, it is an issue of learning mode because otherwise, we might expect a better score from the top performers on either or test. In the event, the top performing students all had their scores lowered when taking the medical tests. Looking at the top performing and the least performing student groups we can see that the top group's performance, regarding sheer improvement, on unknown medical words was close to that of the lower group, i.e. little to no improvement was detectable.

Again, looking at chart 1., it can be seen that the spread for the average score between the top performing to the least performing is 4.67 points or 48%, comparing the general to the medical registries. Although this is certainly a large figure, the pattern for their performance was similar, they both dropped in the score for the medical tests and remained vertically unchanged for week six, nine and twelve. Such a repetitive pattern further suggests that computer-based learning provides little to no appreciable benefit to the students.

Student	Week												Total	Average
	1	2	3	4	5	6	7	8	9	10	11	12		
1	10	10	8	10	10	8	10	10	9	9	10	9	113	9.416666667
2	10	10	7	9	10	8	10	10	7	7	10	8	106	8.833333333
3	9	10	6	10	8	7	10	10	8	8	10	8	104	8.666666667
4	8	9	7	9	10	6	9	10	7	10	8	9	102	8.5
5	9	10	6	10	10	6	10	10	7	10	10	7	105	8.75
6	10	10	4	10	10	5	10	9	6	9	10	6	99	8.25
7	9	9	7	8	9	7	9	8	8	8	8	5	95	7.916666667
8	7	8	4	9	9	5	8	9	7	9	10	7	92	7.666666667
9	8	9	7	8	9	7	8	9	5	8	9	4	91	7.583333333
10	8	9	5	9	9	4	8	8	5	8	8	7	88	7.333333333
11	7	9	7	8	9	6	9	9	4	8	9	6	91	7.583333333
12	8	8	5	7	7	7	7	8	6	9	7	5	84	7
13	7	7	5	7	6	5	6	6	6	7	8	5	75	6.25
14	8	9	6	8	6	5	7	6	5	6	8	7	81	6.75
15	7	7	4	6	6	5	5	7	4	7	8	5	71	5.916666667
16	8	8	7	5	6	6	6	7	5	7	7	6	78	6.5

17	6	8	5	6	7	5	5	6	6	7	6	5	72	6
18	7	6	6	7	6	5	7	7	3	8	7	4	73	6.083333333
19	7	7	3	7	5	6	5	7	4	6	7	5	69	5.75
20	7	5	5	6	5	3	4	6	4	6	7	4	62	5.166666667
21	8	6	4	8	5	5	5	6	5	7	6	3	68	5.666666667
22	6	5	4	7	7	6	6	5	4	6	6	4	66	5.5
23	7	5	4	6	6	4	7	6	6	6	5	4	66	5.5
24	6	7	5	7	6	4	5	6	5	7	6	5	69	5.75
25	6	5	4	5	5	5	7	4	3	7	5	4	60	5
26	6	6	3	7	5	5	7	6	4	6	7	5	67	5.583333333
27	7	6	4	5	4	5	6	6	3	6	6	3	61	5.083333333
28	7	7	4	4	5	4	7	4	4	5	7	3	61	5.083333333
29	6	5	5	6	4	3	8	5	6	7	5	5	65	5.416666667
30	6	5	3	5	5	5	7	6	4	7	6	4	63	5.25
31	5	4	4	7	4	4	5	6	5	5	3	6	58	4.833333333
32	7	5	5	6	4	4	7	6	3	6	5	4	62	5.166666667
33	4	6	4	6	5	4	7	5	4	5	4	4	58	4.833333333
34	6	6	3	5	3	3	7	4	4	6	5	3	55	4.583333333

35	7	5	5	6	4	4	6	6	5	5	6	3	62	5.166666667
36	5	3	4	5	4	5	5	4	3	7	5	3	53	4.416666667
37	6	4	4	6	5	3	5	4	3	6	5	4	55	4.583333333

Chart 1: Individual student scores and averages ● Medical Register ● General Register

Even week numbers = computer-based learning / Odd week numbers = paper-based learning

In comparison, the data shows an increase in performance for general registry with computer-based learning over paper-based learning. This can be explained because these are words that the students should have experience of from junior, and so their familiarity is high and cognitive load is low, benefiting learning. As a result, in this context, computer-based learning is more effective with familiar content, but the reverse has been seen to be true (see above paragraph) with unfamiliar content, mirroring the findings of Wallace and Clariana (2000). Thus, it has been demonstrated that computer-based learning is more beneficial for high-achieving students on familiar content, but low-achieving students receive fewer benefits from this mode of learning.

As a side note, figure. 1 (see above) shows that for week five the students' scores dropped compared to the average score for the general register. The author believes that this was caused by an increase in the student's anxiety, as this week coincided with their mid-term examinations for other subjects and not due to an increase of difficulty in the vocabulary. The data was included nonetheless because although it was a noticeable drop in performance, it still followed the trend for general registry as compared to medical.

Conclusion

In the past, the adoption of computer-based learning has been pursued with great interest because it was thought that computer-based learning would significantly increase all student's performance. However, the data from this study suggests that such a level of interest may be questionable. Although it would appear that higher-achieving students can receive some of the benefits from this mode of learning, weaker and low-achieving students may not be able to receive the same level of benefits. In fact, the latter group may be better served with traditional paper-based assessments. One possible reason for this may be that computer-based learning actually increases

student's cognitive workload, and at a time when they are learning unfamiliar content, this increase no matter how small may be enough to negatively affect their performance. Unfortunately, relatively little research has been conducted on how computer-based learning and testing affects cognitive workload, and as such, this is the first recommendation for further research that this study puts forward.

The authors of this study are in agreement with Emerson and MacKay (2011) on the need for institutions of higher education to pay greater caution in the adoption of computer-based learning in their curriculums and that the overriding concern must be how it will benefit student performance. Certainly, computer-based learning can play an important role in benefiting student performance, but what is called for is a better understanding of the interplay between student performance, learner preference and the mode of learning, i.e. computer-based vs. paper-based learning and how these affect performance.

The authors of this study suggest that educators use computer-based testing as they would any other education method, tool or style; begin by researching and assessing the possible benefits of the new approach over the old and what the ultimate benefits for the students may be. Once adopted, they should engage in constant monitoring and review of the new approach and take into account not only raw student performance but opinions too. They should be prepared to abandon the approach if the benefits to performance are not met or do not justify the expense in monetary or human terms. Particularly with the former, institutions and educators should more carefully consider the initial adoption of computer-based technologies as it is extremely difficult to go back on such decisions if they prove unfruitful as compared to traditional learning modes.

Finally, despite the overall goals of this study being met, namely the production of empirical evidence that suggests caution over the unquestioned adoption of computer-based learning at the expense of traditional modes, the authors feel further research needs to be conducted. The study represented only a relatively small sample size of exclusively engineering students, and so it is suggested that a larger sample, across disciplines, be conducted in the future. Also, a study that directly surveyed student learning mode preference, compared against their performance on said modes, could produce interesting results.

References

- Clariana, R. B., & Wallace, P. E. (2002). Paper-based versus computer-based assessment: key factors associated with the test mode effect. *British Journal of Educational Technology* 33 (5), 593–602. <https://doi.org/10.1111/1467-8535.00294>
- Emerson, L. & MacKay, B. (2011). A comparison between paper-based and online learning in higher education. *British Journal of Educational Technology*, 42 (5), pp. 727-735. doi:10.1111/j.1467-8535.2010.01081.x
- Fluck, A., Pullen, D. & Harper, C. (2009). Case study of a computer based examination system. *Australasian Journal of Education Technology*, 25 (4), pp. 509-523. <https://doi.org/10.14742/ajet.1126>
- Gruba, P., & Hinkleman, D. (2012). *Blending technologies in second language classrooms*. New York: Palgrave Macmillan. <https://doi.org/10.1057/9780230356825>
- Hall, G. (2011). *Exploring English Language Teaching: Language In Action*. London: Routledge.
- Kukulska-Hulme, A., & Shield, L. (2008). An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL*, 20(03). <https://doi.org/10.1017/S0958344008000335>
- Leeson, H. V. (2006). The Mode Effect: A Literature Review of Human and Technological Issues in Computerized Testing. *International Journal of Testing*, 6 (1), pp. 1-24. https://doi.org/10.1207/s15327574ijt0601_1
- Levy, M., & Stockwell, G. (2006). *CALL dimensions: Options and issues in computer assisted language learning*. Mahwah, NJ: L. Erlbaum Associates.
- Low, Ee-Ling. (2014). *Pronunciation for English as an International Language: From research to practice*. London; New York: Routledge.
- Noyes, J. M., Garland, K. J., & Robbins, L. (2004). Paper-based versus computer-based assessment: is workload another effect? *British Journal of Education Technology*, 35 (1), pp. 111-113. <https://doi.org/10.1111/j.1467-8535.2004.00373.x>
- Paas, F., Renkle, A., & Sweller, J. 2004. Cognitive Load Theory and Instructional Design: Recent Developments. *Educational Psychologist*, 38 (1), 1–4. https://doi.org/10.1207/S15326985EP3801_1
- Sweller, J. 1988. Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12 (2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Wallace, P. E., & Clariana, R. B. (2000). Achievement predictors for a computer-applications module delivered via the world-wide web. *Journal of Information Systems Education*, 1 (1), pp. 13-18.

Watson, B. (2001). Key factors affecting conceptual gains from CAL. *British Journal of Educational Technology*, 32 (5), pp. 587-593. <https://doi.org/10.1111/1467-8535.00227>