Testing Effect for Visual-Symbolic Material: Enhancing the Learning of Filipino Children of Low Socio-Economic Status in the Public School System

Daniel Wartenweiler
Onesimo Bulilit Foundation and De La Salle University, Manila

Abstract

Providing tests enhances the retention of learned material compared with further study. This relationship has been known as the testing effect. Studies have shown that the testing effect can also be applied to the transfer of learning, i.e. novel demonstrations of the learned material. The author tested this relationship with Filipino Elementary School children (n = 32) of low socio-economic status in an underprivileged school setting. In a with-subjects experimental design, the participants underwent two different study procedures on two sets of visual symbols, including a study-only and a test-study condition. After a one hour retention interval, they took retrieval and transfer final tests on both sets. A significant testing effect (p < .05) was found only on final tests requiring transfer but not on final tests only requiring retrieval. Even though participants in higher Elementary School grades performed better on both final tests, there was no interaction effect of grade level with the testing effect. The present study confirms that the testing effect also applies to the transfer of learning, and that the testing effect is even bigger for transfer than for simple retrieval. The study also suggests that the testing effect applies to children in all grade levels of Elementary School, to children of low socio-economic status in the Philippine public school system, and to visual-symbolic materials.

Keywords: testing effect, retrieval practice, retention, transfer, learning, visual-symbolic memory, education, street children, public school, Philippines

Testing effect

Studies have shown that material is better remembered when the study of the material included at least one test of the material, compared re-studying the material. This so-called testing-effect has been observed for a wide variety of materials, such as word lists (Carpenter & DeLosh, 2006; Shana, & Harold, 2007; Karpike, & Zaromb, 2010), text passages (Roediger & Karpicke, 2006b; Agarwal, Karpicke, Kang, Roediger, & McDermott, 2008; Richland, Kornell, & Kao, 2009), face-name pairings (Carpenter & DeLosh, 2005), multimedia materials (Johnson & Mayer, 2009), facts and concepts (Butler, 2010; Jacoby, Wahlheim, & Coane, 2010), visual-spatial materials (Carpenter & Pashler, 2007; Rohrer, Taylor & Sholar, 2010), foreign language vocabulary (Toppino & Cohen, 2009), and skills learning (Kromann, Jensen, & Ringsted, 2009).

The testing effect has also proven to be robust throughout different kinds of recall tests, such as free-recall tests (Roediger & Karpicke, 2006b), cued recall tests (Carpenter & DeLosh, 2006), multiple choice tests

 ISSN 2094-1420

(McDaniel, Anderson, Derbish, & Morisette, 2007; Odegard, & Koen, 2007), recognition memory and classification (Jacoby, et al., 2010), open and closed book tests (Agarwal, et al., 2008), group quizzes (Cranney, Mihyun, McKinnon, Morris, & Watts, 2009), and tests requiring transfer of learning (Butler, 2010; Rohrer et al., 2010).

Even if retrieval during the test was unsuccessful or occurred before the actual study situation, it still improved retention, as long as actual answers are given after the initial test. In fact, unsuccessful tests may enhance future learning by providing additional challenge to the learner (Richland, Kronell, & Kao, 2009).

Feedback on tests has been shown to facilitate the correction of errors and to improve retention (McDaniel, et al., 2007; Voydanoska, Cranney, & Newell, 2010). On the other hand, a testing effect was still evident with no corrective feedback or additional exposure to the material (Roediger & Kaprike, 2006b; Carpenter & DeLosh, 2005).

Metacognitive awareness of the benefits of a learning strategy involving tests was found to enhance memory retention (Jacoby, et al., 2010). On the other hand, students who re-studied material instead of taking tests showed increased confidence that they would remember the material, but they performed worse than those who took tests (Roediger & Kaprike, 2006a; Agarwal, et al., 2008). This finding shows that they were unaware of the benefits of testing.

Roediger and Kapricke (2006b) found an interaction between the retention interval and the testing effect, whereas for a retention interval of five minutes, further study improved recall; but for a retention interval of two days or one week, testing produced substantially greater retention. Toppino and Cohen (2009) found the same interaction, whereas recall for study was insignificantly better for a two and five minutes retention interval, but recall for testing was significantly better after a two days retention interval. These results suggest that testing improves long term retention while additional study may only be of advantage for very short retention intervals.

The strong benefits of the testing effect has led many researchers to advocate for the more frequent use of tests in the educational setting in all levels, not just to assess learning, but as a strategy to improve learning and enhance retention (e. g. Roediger & Kapricke, 2006a; Agarwal, et al., 2008; Johnson & Mayer, 2009; Rohrer, et al., 2010). Recent research has investigated the application of the testing effect with college students in the real classroom setting, and has demonstrated improved retention through the use of quizzes as a learning strategy (McDaniel, et al., 2007; Cranney, et al., 2009).

Theoretical Accounts of the Testing Effect

The *transfer-appropriate processing* hypothesis predicts that memory performance not only depends on the depth of processing, but also on the similarity of the processes involved in encoding and decoding (Morris, Bransford, & Franks, 1977). According to the transfer-appropriate processing theory, the testing effect can be explained by the similarity between the learning strategy and the manner of the final test. The inclusion of initial tests in the learning strategy is more effective because the initial test provides transfer-appropriate processing and increases the chance of final retrieval. For example in a study conducted by Johnson & Mayer (2009), participants performed better on a delayed retention test after a practice retention test than after a practice transfer test (a test requiring novel demonstration of learning); but the opposite pattern was true on a delayed transfer test. Also, a testing effect was found for both for the practice retention test group and the practice transfer test group versus the restudy groups. The results of this study support the explanation of the testing effect based on transfer-appropriate processing. Other studies however have not supported the conjuncture that the match of initial and final test format leads to better results on the final test. (e.g. Carpenter & DeLosh, 2006). In many testing effect studies however, initial and final test were identical (e.g. Kuo & Hirschmann, 1996; Carpenter & DeLosh, 2005; Roediger & Karpicke, 2006b).

The finding that the retention interval interacts with the testing effect (Roediger & Kaprike, 2006b; Toppino & Cohen, 2009), suggests that different processes are involved in studying and testing. An alternative explanation of the testing effect attributes the testing effect to the retrieval process that happens during testing (Glover & Krug, 1990). Roediger and Karpike (2006a) discussed how retrieval during an initial test might benefit retention. They suggest that retrieval might help to differentiate material well learned from material less well learned, and therefore help to apply encoding strategies to items that most need it. Even if no corrective feedback is provided (Roediger & Kaprike, 2006b; Carpenter & DeLosh, 2005) or if the retrieval during the test was unsuccessful (Richland, Kronell, & Kao, 2009), tests still enhance memory retention. Roediger and Kaprike (2006a) referred to this as the direct benefits of testing. In other words: the retrieval processes involved in testing alone and even without correction of errors enhance memory of the tested material.

The *elaborative retrieval hypothesis* explains this finding through "elaborative information related to the target response" that is activated during the retrieval process and thereby increases "the chances that activation of any of this information will facilitate later retrieval of the

target" (Carpenter, 2009). Carpenter and DeLosh (2006) investigated the elaborative retrieval hypothesis in contrast with transfer-appropriate processing. They used wordlists as study material and manipulated the format of initial and final test in a factorial design, using both initial and final tests that required recognition, cued recall and free recall. Retention was best for final tests when the initial test was a free recall, regardless of the type of the final test. The finding suggests, contrary to transfer-appropriate processing view, that the testing effect is not based on the similarity of the initial and final test, but on the quality of the retrieval processes involved in the initial test, predicting that performance on a final test is greater if the initial test required greater retrieval difficulty. Carpenter (2009) in a follow-up study used cue-target word pairs that were strongly associated and cue-target word pairs that were weakly associated with both a cue-target recall test and a restudy condition. A final test requiring free recall of the targets showed that tested items were retained better than studied items, and targets recalled from weak cues were retained best over time. Strong cues produced quick and easy answers on an initial test as evidenced by the faster and more accurate recall on the same test, but weak cues needed longer processing and seemed to have produced more elaborative information on the target, leading to better recall over time.

Other studies have also shown that greater difficulty on an initial test produces better retention. For example, Kaprike and Zaromb (2010) found that intentional retrieval produced better retention than incidental retrieval. Also, learning procedures including production tests (short answer or essay) have been found to be better in improving retention than cognition multiple-choice tests (Mc Daniel, et al. 2007). These studies provide further evidence that indeed elaborative retrieval processes play an important role in the testing effect.

Studies on the Testing Effect with a Final Test Requiring Transfer

Rohrer et al. (2010) raised the important question if the testing effect diminishes if the final test requires transfer of learning. Transfer is understood as a novel demonstration of learning where material learned does not only need to be recalled in the exact same form, but has to be transformed into different forms and applied to new situations. Transfer is sometimes said to be the aim of learning. According to the transfer-appropriate hypothesis, transfer on a final test should reduce the size of the testing effect, since initial and final test are dissimilar. If the testing effect really diminishes in size when applied to the transfer of learning; that would mean that the testing effect does not really have much practical application in the educational setting. In other words: if a learning strategy that involves

testing only improves performance on exactly the same test than the initial test, then it could only be applied to settings were format and content of the final test are already known.

Only a few studies have investigated the testing effect on the transfer of learning. McDaniel, et al. (2007) conducted an experimental study with college students in the classroom setting. They used paired questions based on the reading assignments from an undergraduate textbook. The questions represented a forward and reverse association of the same fact (i.e. A => B vs. B => A). One of the paired questions appeared on the initial test, and the other on the final test. Initial tests enhanced learning compared to additional reading, and the testing effect was significant even with a final test requiring transfer.

Johnson and Mayer (2009), used a narrated animation about lighting that was either re-watched by the participants, or tested on a practice retention test or a practice transfer test. The participants that were tested performed better on a final test than those who re-watched the animation, and that held true also for final test questions requiring transfer.

Butler (2010) used different passages about various topics and administered repeated study and testing conditions on the material. The repeated testing conditions included the same questions, new inferential questions based on the same knowledge domain, and inferential questions based on different knowledge domains. Repeated testing produced better results both on retention and transfer final tests.

Rohrer, et al. (2010) used maps in two different learning procedures including a study-only condition and a study-test condition. Testing effects were highly significant on both retention and transfer final tests, where as the testing effect was even slightly bigger for transfer.

Methodological Problems in Testing Effect Studies

Carrier and Pashler (1992) pointed out methodological problems in testing effect studies. Using the same exposure time for studied and tested word pairs, they reported solid testing effects and proved that the testing effect is not confounded through further study opportunities provided during the corrective feedback of initial tests. Kuo and Hirshman (1996) listed three methodological problems that could confound the testing effect. First, low retrieval rates on the initial test may produce an underestimate of the testing effect. Second, presentation conditions of re-study and test might be different. Specifically, during test conditions, sometimes participants write down tested items and are able to see all the previously administered items. The testing effect might therefore be confounded through simultaneous viewing. Third, if a sequence of test and study conditions is being administered, previously

tested items might be re-studied, and the testing effect might be due to re-study rather than due to different processes involved in testing and studying. Kuo and Hirshman (1996) used three word lists that were administered in different sequences of study and test conditions. The same output conditions for test (T) and study (S) procedures were used, and the three word lists produced high retrieval rates on test conditions. A study-test sequence of STTTT produced better performance on a final retrieval test than a SSSSS condition, demonstrating that re-studying of previously tested material was not confounding the testing effect. The authors argue that the processes underlying test and study differ and they suggest that item-specific factors contribute to the testing effect.

The present research has adopted the methodology of Rohrer, et al. (2010), which addresses the three methodological problems discussed by Kuo and Hirshman (1996). Low retrieval rates on the initial test have been counteracted by providing immediate corrective feedback. The feedback was however so short that it provided very limited opportunity for further study. Output conditions for both study and test cycles were exactly the same, and simultaneous viewing of previous and present items was prevented because only sequential numbers were written down by the participants. The methodology made use of a re-study procedure of five cycles and a test procedure of five cycles, following a format of SSSSSS and STTTTT. Both test and study procedures had exactly the same total exposure time for each item.

Rationale for the Present Research

There are only a few studies available so far that investigate the testing effect on the transfer of learning, and all those studies have been done in the US context. Since there have been no studies reported in the Filipino context and with children from low socio-economic status deprived of educational opportunities, the author wanted to test if the testing effect also applies to children with limited academic stimulation in the Philippines.

The studies previously reported provide evidence that the testing effect does also apply to transfer, even though only Rohrer, et al. (2010) so far have provided results were the testing effect was greater on a transfer final test than on a retention final test. The author of the present study tested this conjecture by reduplicating the experiment reported by Rohrer, et al. (2010). Instead of using maps as the learning material, the author used visual symbols with an assigned meaning, making use of visual-symbolic and verbal memory in the assumption that the combination of different learning styles will benefit underprivileged children. Since the author is working among street children in Manila with an NGO providing residential care and community based interventions such as educational assistance and non-

formal education, it is of his personal interest to work towards improving the educational opportunities and learning strategies of underprivileged and academically delayed children.

The author tested the hypotheses that there is a significant testing effect on both a transfer and a standard final test with underprivileged Filipino children in the public education system for visual-symbolic memory learned with a test-study learning strategy compared to material learned with a study-only learning strategy. Furthermore, the author tested the hypothesis that the testing effect on transfer final test is greater than the testing effect on a standard final test, using the same visual-symbolic materials.

Method

Overview

The experiment investigated the testing-effect of visual-symbolic material both for a standard and a transfer final test with Filipino Elementary School children of low socio-economic status in the public education system. The students memorized visual symbols that represented nouns in the Philippine National Language (Tagalog). The so-called standard or retrieval final test required participants to perform the same test they had practiced during the learning session: assign the right noun to the corresponding symbol. On the transfer final test, they had to answer ten questions involving transfer of the learned symbol meanings.

Participants

Both learning and test sessions were completed by 32 Filipino Elementary School students of low socio-economic status who were enrolled in the public school system. The participants were 8-14 years old (M = 10.6, SD = 1.7) and enrolled in Grade 2-6 (M = 3.7, SD = 1.4). 17 were in lower elementary level (Grade 2-3) and 15 in higher elementary (Grade 4-6). Only 12 (37.5%) children were in the grade that corresponds with their age, while 20 (62.5%) were advanced in age. The mean years participants were advanced in age was 0.97 (SD = 1.03). 19 were female (59.4%) and 13 male (40.6%). All participants were clients of the Educational Assistance Program or the Residential Shelter of Onesimo Bulilit Foundation. They qualified for these programs due to their low socio-economic status. They had been living on the streets, under bridges or in make-shift shelters along the sidewalks, and were deprived of educational opportunities. The children in the Residential Shelter either came from abusive family situations and / or their

parents were in jail. The fact that 62.5% of the children were over age compared to their grade level shows that that schooling has been a low priority for their caregivers.

The public school which the participants attended has class sizes of 50-60 students, and the students go to class either on a morning schedule from 6am to 11:30am or on an afternoon schedule from 12pm to 5:30pm. Language of instruction is mostly Tagalog, but the nationwide exams have written instructions in English, except for the Filipino (Tagalog) subject. Children for underprivileged backgrounds are not exposed to English in their homes or communities, though English is a subject being taught. Students are mainly taught through rote learning and memorization. There is a limited emphasis on comprehension, critical thinking, and metacognition.

Materials

Two sets of 10 visual symbols each with an assigned meaning of a two-syllable noun in Filipino (Tagalog) were used for the learning and test procedures. The nouns were chosen from a reading book for Grade 1 that had been studied by all students. The nouns could be easily read with a Grade 2 reading level and they were familiar to all participants due to daily usage. Some of the nouns were objects of daily life, some food items, and some were animals. The symbols were taken from Wingdings and Wingdings2 fonts. Both sets consisted of a few symbols each that were similar in shape and design, in order to make memorization more challenging. The meanings were randomly assigned to the symbols. An additional sample set of 5 visual symbols was used during tutorials (See Figure 1).

Procedure

The 32 participants were assigned to four groups by matching them according to their grade level. Each group then underwent a tutorial procedure, followed by a test-study (TS) and a study only (SO) learning procedure on one set each. The materials were presented to each group on a 15-inch computer screen via timed power point presentation.

Both the order of the TS and SO procedures and the pairing of the procedures and set of symbols were counterbalanced so that each of the matched groups was randomly assigned to one of the four different schedules. Group 1 underwent first a TS procedure on set 1 followed by a SO procedure on set 2. Group 2 underwent first a SO procedure on set 1 followed by a TS procedure and set 2. Group 3 did TS first on set 2 and than SO on set 1, while Group 4 did SO on set 2 first followed by TS on set 1. The counterbalancing ruled out that the variable of two different sets used and the sequence of the

learning procedure confounded the testing effect. Group 1 and 3 underwent their learning procedure simultaneously in two separate rooms. After the learning procedure, both groups watched for one hour a part of the movie *Narnia*, which served as distraction before the final retrieval tests. While Group 1 and 3 watched the movie, Group 2 and 4 underwent their learning procedures. After their learning procedure, they also watched for one hour a part of the movie *Narnia.* After the one-hour retention interval, Group 1 and 3 took their final tests first, followed by the final tests Group 2 and 4 also after on hour of movie watching. Both transfer final tests were administered first, followed by both standard final tests, in order to rule out that recall during the supposedly easier standard final test would improve performance on the transfer final test, and therefore confounding the testing effect on the transfer final test. Half of the participants (group 1 and 3) were administered the final test for the set learned with the TS procedure first, in order to rule out order effects of the set studied first.

Tutorial. Before the learning procedure, each group first underwent a tutorial procedure with the sample set familiarizing them with both learning procedures and with a sample final test. They underwent an initial exposure cycle, where each of the five symbols together with its meaning was shown on the screen (black font on white background) for 6 seconds. A sequential number was also shown at the top of the screen (see figure 2). During the initial exposure cycle, the participants were asked to just watch and study the symbols with their meanings. Then the participants cycled through four learning cycles, two according to the SO procedure and two according to the TS procedure (see below). Once they finished the learning procedures, they underwent a sample transfer final test that consisted of three questions.

SO Procedure. The participants were given a worksheet that five tables with all the ten symbols listed and blank spaces behind them, one table for each learning cycle. They then underwent an initial exposure cycle which was the same than during tutorials. After initial exposure, the participants cycled through five learning cycles. Each symbol was now shown for 10 seconds together with its meaning. A sequential number was also shown at the top portion of power point slide. The participants were instructed to write the sequential number shown on the blank space behind the correct symbol on their worksheet. The sequence of the 5 cycles was varied systematically across cycles so that no two symbols appeared consecutively in more than one cycle. Total exposure time for each symbol therefore was 6 seconds initial exposure plus 50 seconds re-study, a total of 56 seconds.

TS Procedure. The TS procedure was basically the same as the SO Procedure. There was one initial exposure cycle followed by five learning cycles. The participants also used the same format of worksheet to write the sequential numbers. During the learning cycles, however, for the first 7 seconds only the meaning was shown without the symbol and the participants were instructed to write the sequential number shown on the blank space behind the correct symbol on their worksheet. Immediately afterwards, the correct symbol was faded in during 0.5 seconds, and was shown for 2.5 seconds, to provide immediate correction of the answer given. Total time spent on each symbol in each cycle was therefore 10 seconds, exactly the same than in the SO procedure.

Standard final test. The participants were provided with a list of the symbols and a numbered list of the meanings, and they had to assign the correct number to each of the 10 symbols. The standard final test therefore followed exactly the same format than the learning procedure, were assignment of the correct meaning to the symbol was required. The accuracy of the answers was calculated in percent.

Transfer final test. The participants were asked a total of 10 questions requiring transfer of learning, one for each symbol. The transfer questions were inherently more difficult, since they required the participants to either relate newly learned material to previous knowledge, or to recall and produce material through drawing or writing that was previously learned through assignment. Three questions were multiple choice questions, for example: *Bilugan ang hayop* (cycle the animal): ⬟ ◤ ▼. Seven were production questions that required either recall of the meaning or recall of the symbol. Examples are:

Iguhit ang symbolo ng (draw the symbol of): bata _____.

Isulat ang ibig sabihin ng (write the meaning of): ➔ _____.

The accuracy of the answers was calculated in percent.

Analysis

For the data analysis, *t*-tests were run for the two learning strategies (TS and SO) for both final tests (transfer and standard). Additionally, Cohen's *d* was calculated for the testing effect of both final tests. *T*-tests were run to rule out order effects between the set studied first and set studied last. A 2x2 ANOVA was run to rule out that grade level (lower elementary vs. higher elementary) interacted with the testing effect on each final test.

Results

There was a mean difference between SO and TS learning strategy on both the standard final test (M_{SO} = 57.8%, M_{TS} = 64.1%) and the transfer final test (M_{SO} = 54.1%, M_{TS} = 60.0%). The testing effect however was only significant for the transfer final test, $t(31) = 2.07$, $p < .05$, but not for the standard final test. For both final tests, the effect size was small, $d = 0.21$ (transfer) and $d = 0.18$ (standard).

Order effects of the set first studied and the set last studied were insignificant for both transfer and standard final tests.

Even though there was a significant main effect for grade level (lower elementary vs. higher elementary) on both transfer final test, $F(1,30) = 5.30$, $p < .05$, and standard final test, $F(1,30) = 5.53$, $p < .05$, there was no significant interaction effect of grade level and learning procedure (SO vs. TS).

Discussion

A test-study learning condition only significantly improved the accuracy on a final test requiring transfer of the memorized visual-symbolic material, but not on a standard retrieval final test. The testing effect on the transfer final test was slightly bigger than the testing effect on the standard final test; however, both effect sizes were small. Since order effects were insignificant and there was no interaction of grade level with the testing effect, it can be concluded that the test-study learning condition improved performance on the transfer final test, but not on the standard final test. The testing effect was therefore greater on the transfer final test than on the standard final test.

The study of Rohrer et al. (2010), in contrast, reported significant testing effects for both final tests on both replications of the experiment on a significance level of *alpha* < .01, while testing effects for transfer final tests were greater ($d = .76$ and $d = .57$) than testing effects for standard final tests ($d = .64$ and $d = .54$). Nevertheless, the study confirms the finding the testing effect does not diminish for final tests requiring transfer, but is actually larger than on retrieval final tests.

Rohrer et al. (2010) showed that the testing effect can be applied to children in Elementary School. The present study suggests that the testing effect applies to children in all grade levels of Elementary School, to children of low socio-economic status in the Philippine public school system, and to visual-symbolic materials.

There is a significant effect for grade level on both final tests, showing that children in the higher grade level performed significantly better on both

final tests than children in the lower grade level. Further studies limiting the variance in the grade level may produce data with a smaller variance, therefore enhancing the testing effect.

Further factors might also have contributed to a smaller testing effect in the current study.

In further studies, the test-study learning procedure should be improved. During the test-study condition, the instruction was given to the participants to answer on their worksheet while only the meaning was shown, and afterwards check their answer when the symbol appeared. Nevertheless, the examiner observed some of the participants waiting for the solution until they wrote their answer. It is possible that they didn't really try to retrieve the right answer during the test phase, compromising the quality of the test-study condition. The participants may not have had the metacognitive awareness of the benefits of the test-study condition, which has been proven to enhance the testing effect. An improvement in the test-study condition could also be to alternate test and study cycles, without giving correct answers during test cycles, but providing feedback through re-study after testing. Also, providing further information to the participants about test-based learning strategies could improve their metacognitive awareness. Further studies may also consider including metacognitive measures (compare with Jacoby, et al., 2010).

Rohrer et al. (2010) used retention intervals of one day, while the present study used an interval of one hour. Previous studies have shown that the testing effect increases with increased retention interval. Future studies may either extend the retention interval or include groups with different intervals.

It might also be possible that the testing effect is greater for children in a private school setting who are used to enhanced methods of learning compared to children of low socio-economic status in the public school system. One factor might be that the public school system does not raise the awareness of the benefits of test-enhanced learning strategies, but mainly emphasizes learning based on repetition and memorization. Further research will be needed to test this hypothesis by administering similar experiments to children of different socio-economic statuses in public and in private schools.

The present study was able to demonstrate the benefits of test-enhanced learning strategies applied to the Philippine school setting. Such strategies remain underutilized, especially in the public school system. Educators should make regular use of tests as a learning strategy, and not just to assess learning. They should also work towards educating children on the benefits of tests, and how they can improve their learning by making use of tests and feedback.

References

Agarwal, P., Karpicke, J., Kang, S., Roediger III, H., & McDermott, K. (2008). Examining the testing effect with open- and closed-book tests. *Applied Cognitive Psychology*, *22*(7), 861-876. doi:10.1002/acp.1391

Butler, A. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(5), 1118-1133. doi:10.1037/a0019902

Carpenter, S. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(6), 1563-1569. doi:10.1037/a0017021

Carpenter, S., & Pashler, H. (2007). Testing beyond words: Using tests to enhance visuospatial map learning. *Psychonomic Bulletin & Review*, *14*(3), 474-478.

Carpenter, S., & DeLosh, E. (2006). Impoverished cue support enhances subsequent retention: Support for the elaborative retrieval explanation of the testing effect. *Memory & Cognition*, *34*(2), 268-276.

Carpenter, S., & DeLosh, E. (2005). Application of the testing and spacing effects to name learning. *Applied Cognitive Psychology*, *19*(5), 619-636. doi:10.1002/acp.1101

Carrier, M., & Pashler, H. (1992). The influence of retrieval on retention. Memory and Cognition, 20 (6), 633-642.

Chan, J., & McDermott, K. (2007). The Testing Effect in Recognition Memory: A Dual Process Account. *Journal of Experimental Psychology / Learning, Memory & Cognition*, *33*(2), 431-437. doi:10.037/0278-7393.33.2.431

Cranney, J., Mihyun, A., McKinnon, R., Morris, S., & Watts, K. (2009). The testing effect, collaborative learning, and retrieval-induced facilitation in a classroom setting. *European Journal of Cognitive Psychology*, *21*(6), 919-940. doi:10.1080/09541440802413505

Glover, J., & Krug, D. (1990). The `testing' effect and restricted retrieval rehearsal. *Psychological Record*, *40*(2), 215.

Jacoby, L., Wahlheim, C., & Coane, J. (2010). Test-enhanced learning of natural concepts: Effects on recognition memory, classification, and metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(6), 1441-1451. doi:10.1037/a0020636

Johnson, C., & Mayer, R. (2009). A testing effect with multimedia learning. *Journal of Educational Psychology*, *101*(3), 621-629. doi:10.1037/a0015183

Karpicke, J., & Zaromb, F. (2010). Retrieval mode distinguishes the testing effect from the generation effect. *Journal of Memory & Language*, *62*(3), 227-239. doi:10.1016/j.jml.2009.11.010

Kromann, C., Jensen, M., & Ringsted, C. (2009). The effect of testing on skills learning. *Medical Education*, *43*(1), 21-27. doi:10.1111/j.1365-2923.2008.03245.x

Kuo, T., & Hirshman, E. (1996). Investigations of the testing effect. *American Journal of Psychology*, *109*(3), 451.

McDaniel, M., Anderson, J., Derbish, M., & Morrisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology*, *19*(4/5), 494-513. doi:10.1080/09541440701326154

McDaniel, M., Roediger, H., & McDermott, K. (2007). Generalizing test-enhanced learning from the laboratory to the classroom. *Psychonomic Bulletin & Review*, *14*(2), 200-206.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16,* 519-533.

Odegard, T., & Koen, J. (2007). "None of the above" as a correct and incorrect alternative on a multiple-choice test: Implications for the testing effect. *Memory*, *15*(8), 873-885. doi:10.1080/09658210701746621

Richland, L., Kornell, N., & Kao, L. (2009). The pretesting effect: Do unsuccessful retrieval attempts enhance learning?. *Journal of Experimental Psychology: Applied*, *15*(3), 243-257. doi:10.1037/a0016496

Roediger III, H., & Karpicke, J. (2006a). The Power of Testing Memory Basic Research and Implications for Educational Practice. *Perspectives on Psychological Science (Wiley-Blackwell)*, *1*(3), 181-210. doi:10.1111/j.1745-6916.2006.00012.x

Roediger III, H., & Karpicke, J. (2006b). Test-Enhanced Learning. *Psychological Science (Wiley-Blackwell)*, *17*(3), 249-255. doi:10.1111/j.1467-9280.2006.01693.x

Rohrer, D., Taylor, K., & Sholar, B. (2010). Tests enhance the transfer of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(1), 233-239. doi:10.1037/a0017678

Son, L. (2007). Introduction: A metacognition bridge. *European Journal of Cognitive Psychology*, *19*(4/5), 481-493. doi:10.1080/09541440701352317.

Stokes, P., Lai, B., Holtz, D., Rigsbee, E., & Cherrick, D. (2008). Effects of practice on variability, effects of variability on transfer. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(3), 640-659. doi:10.1037/0096-1523.34.3.640

Toppino, T., & Cohen, M. (2009). The testing effect and the retention interval: Questions and answers. *Experimental Psychology*, *56*(4), 252-257. doi:10.1027/1618-3169.56.4.252

Vojdanoska, M., Cranney, J., & Newell, B. (2010). The testing effect: The role of feedback and collaboration in a tertiary classroom setting. *Applied Cognitive Psychology*, *24*(8), 1183-1195. doi:10.1002/acp.1630

Author Note

Daniel P. Wartenweiler, Department of Psychology, De La Salle University, Manila.

Correspondence concerning this article should be addressed to Daniel Wartenweiler, Onesimo Bulilit Foundation, Inc., 390 Fraternal Street, Quiapo, Manila, Philippines. E-mail: d.wartenweiler@onesimo.ch

Appendix A

Set 1

O	suka	*vinegar*
□	pera	*money*
▼	aso	*dog*
↑	lobo	*balloon*
→	tabo	*scoop*

↙	susi	*key*
⬟	lapis	*pencil*
+	bata	*child*
◤	baso	*glass*
◈	ube	*yam*

Set 2

♦	kaha	*box*
☺	sopas	*soup*
✋	tela	*cloth*
☞	ahas	*snake*
✹	poso	*pump*

♥	sili	*chilli*
♣	pito	*whistle*
♠	kuko	*nail*
♦	pusa	*cat*
★	gatas	*milk*

Tutorial Set

➢	bato	*stone*
☑	ibon	*bird*
☾	mesa	*table*
⊖	bahay	*house*
✕	tali	*string*

Figure 1: The twenty visual symbols used for the experiment (set 1 and 2) and for the tutorial procedure (tutorial set) with their assigned meaning in Filipino (Tagalog) and the English translation.

Appendix B

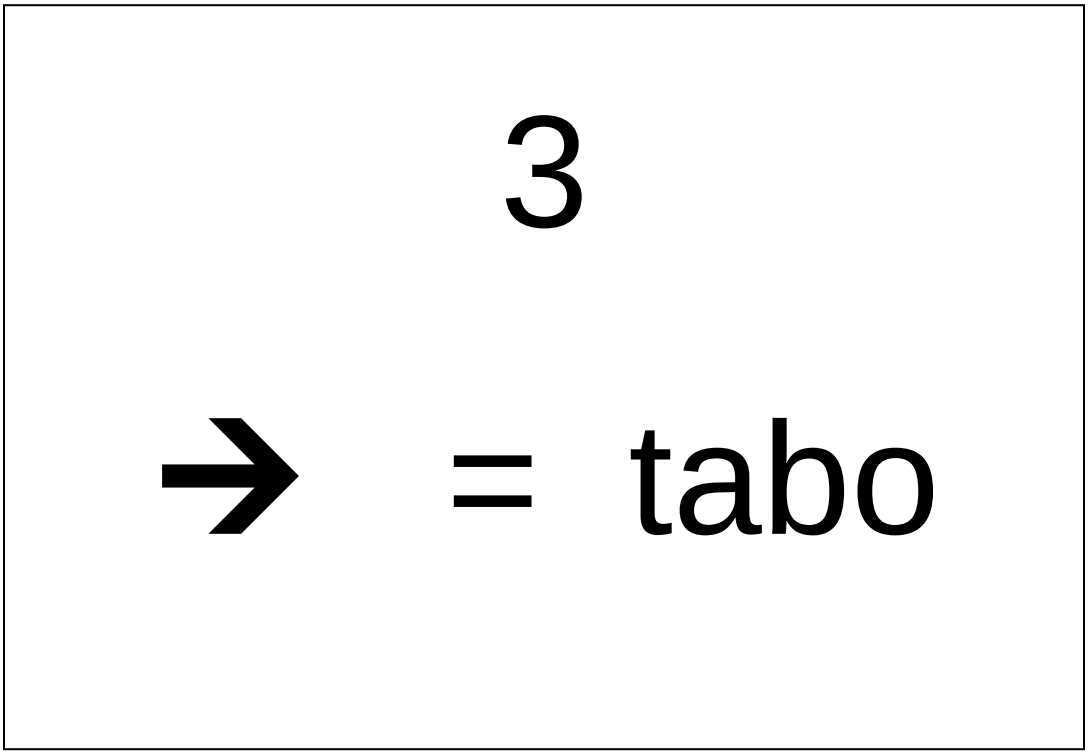

Figure 2: Screen display during initial exposure and SO learning procedure. A sequential number was shown on the top of the screen, and the symbol together with its meaning was shown below. Exposure lasted for 10 seconds.

Appendix C

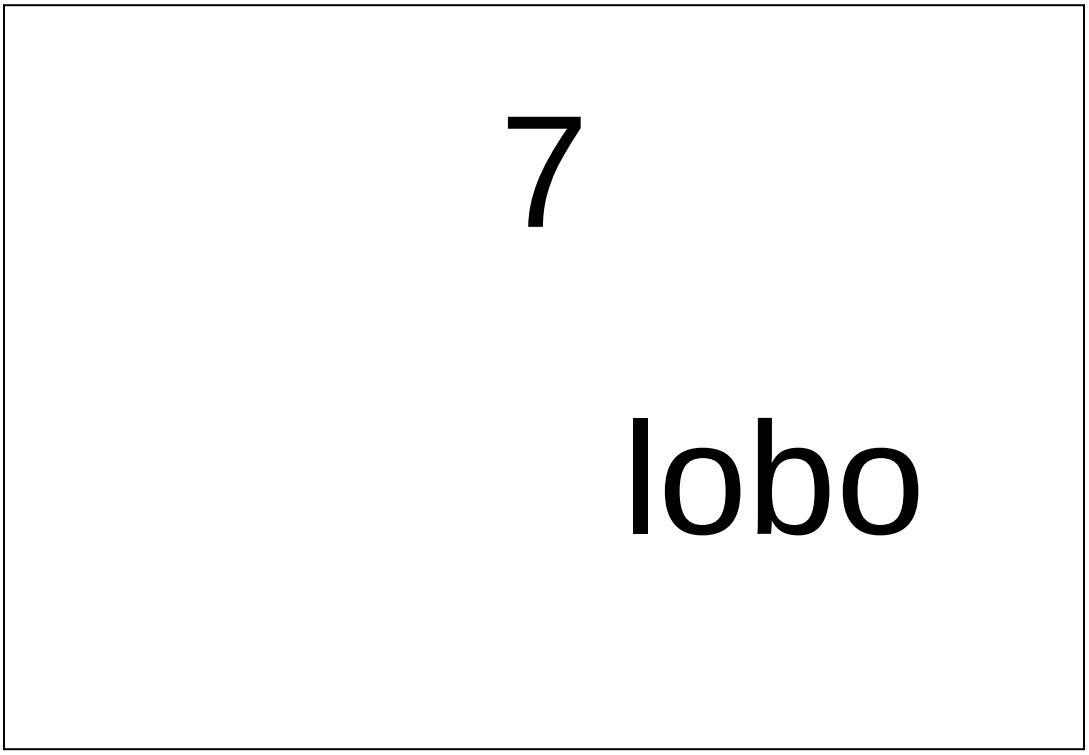

Figure 3: Screen display during TS learning procedure. A sequential number was shown on the top of the screen, and the meaning was shown at the bottom. The correct symbol was faded in after 7 seconds. Total exposure lasted for 10 seconds.

Appendix D

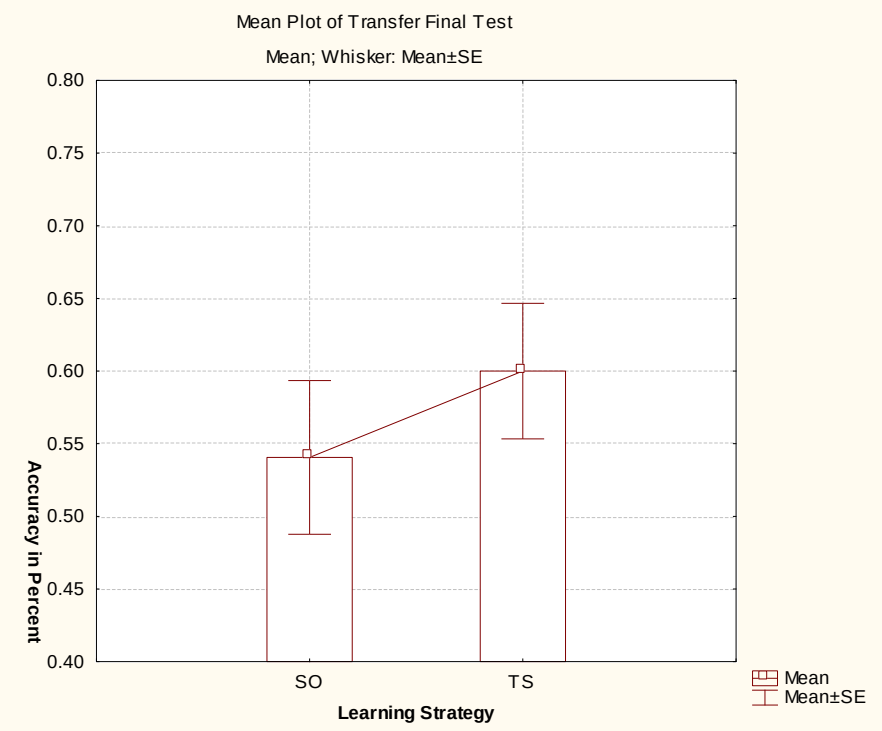

Figure 4. Mean differences in percent on the transfer final test for study-only (SO) and test-study (TS) learning strategy

Appendix E

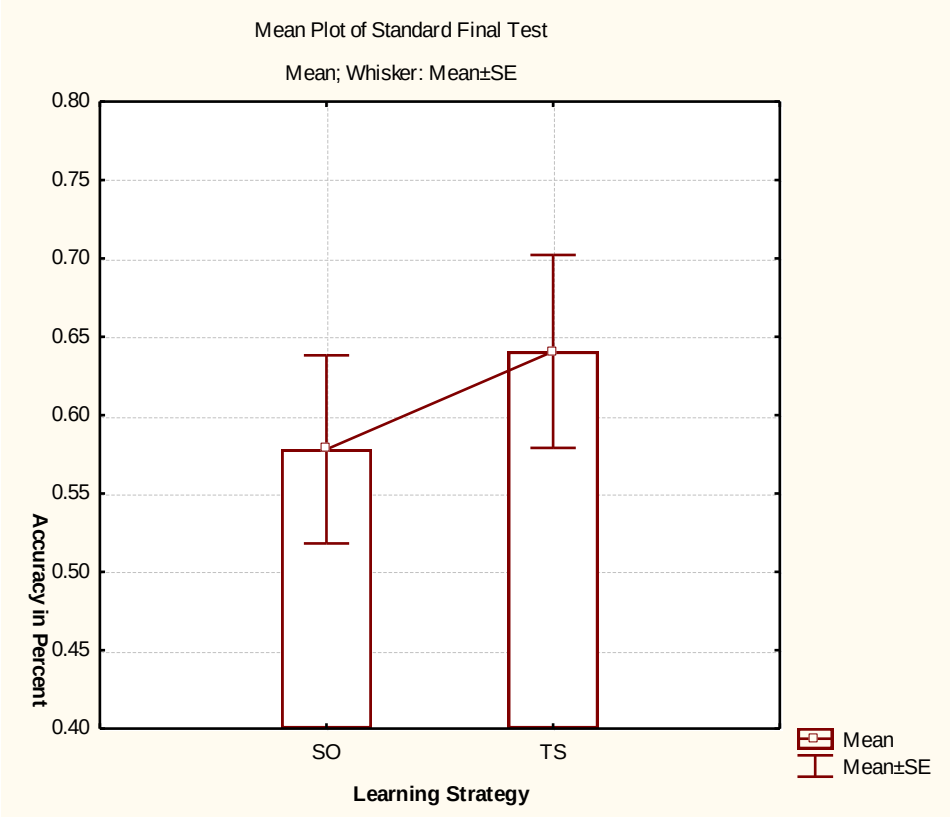

Figure 5. Mean differences in percent on the standard final test for study-only (SO) and test-study (TS) learning strategy.