

REVIEW ARTICLE



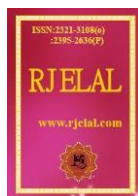
INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA
2395-2636 (Print):2321-3108 (online)

A REVIEW OF RESEARCH ON RETRIEVAL AND TRIAL-AND-ERROR PRACTICES IN SECOND LANGUAGE ACQUISITION

SONG ZIMIAO

Ocean University of China

E-mail: oaimizgnos@163.com



Article info

Article Received: 12/10/2022

Article Accepted: 15/11/2022

Published online:20/11/2022

DOI: [10.33329/rjelal.10.4.82](https://doi.org/10.33329/rjelal.10.4.82)

Abstract

in order to tackle second language (L2) learners' inadequate exposure to nativelike input and facilitate their memorization of L2 knowledge, various practices have been applied in the English-as-a-foreign-language (EFL) classroom. In recent years, there has been a growth in research on the effectiveness of L2 practices, and mixed results have been found. There is an agreement that multiple variables, including practice conditions, linguistic difficulty, and individual differences, are intertwined to affect the learning gains through practice. The present article focuses on two L2 practice conditions that differ in the placement of the definition: retrieval and trial-and-error practices. The aims are to review existing research on retrieval and trial-and-error practices, summarize variables that have been demonstrated to affect the effectiveness of these two practices, systematically explain the mixed results with reference to the desirable difficulty framework, and give suggestions for future studies. Both practices have been demonstrated to be effective in helping to enhance memory of L2 items, while the results vary in different studies with different learning situations. Three important variables that have been demonstrated to play a role in retrieval and trial-and-error practices in prior studies are reviewed here: practice pattern, item number, and learner proficiency. The findings support the conclusion that the practice is optimal when its difficulty matches linguistic and learner-related difficulties. The research in this respect might inform pedagogy in textbook design and teaching procedure.

Keywords: L2 practices; retrieval; trial-and-error; second language acquisition

Introduction

There is a common agreement that the lack of a massive amount of exposure to nativelike input is one of the reasons that L2 learners acquire a second language at a slow speed. Most L2 learners are exposed to a foreign language mainly in the classroom and have few opportunities to use the language outside of the classroom. Various practices

have been applied in the EFL classroom in order to help them enhance the memorization of L2 knowledge efficiently and effectively. Although practices are common in both classrooms and textbooks, research on the effectiveness of practices in helping L2 development is still scarce. Practices are designed to increase the exposure and, ultimately, to establish fluency in the sense of the

smooth operation of psycholinguistic processes, which is necessary for building up skills in language learning. L2 practice is defined as “specific activities in the second language, engaged in systematically, deliberately, with the goal of developing knowledge of and skills in the second language” (DeKeyser, 2007).

The present article focuses on two common practices: retrieval and trial-and-error practices. The distinction between these two practices lies in the presentation order of practices and definition. Retrieval practices present the definition of targets before L2 learners encounter them in reading. Since the learners have already learned the definition, they only need to retrieve it from memory in the following practices, so the chances of making an error are relatively low. That is, the retrieval practices tend to be errorless. In the field of memory, retrieval practices have been demonstrated to be effective in improving one’s memory for material (see Roediger and Karpicke 2006, for a review). The benefit of retrieval is also known as the testing effect in cognitive psychology. The logic behind the testing effect is that learners might invest substantial cognitive effort during retrieving knowledge, which elicits desirable difficulty and entrenches the knowledge. The trial-and-error practices require the learners to make a guess about the definition of targets at the very beginning before it is clearly presented. Since the learners have no prior knowledge that is directly related to targets and have to make a guess, their rate of errors is much higher than that of the retrieval practices. Therefore, trial-and-error practices are generally error-prone. It is assumed that this type of practice might draw learners’ attention to the target items and pique learners’ curiosity, thus enhancing subsequent learning. In essence, the practices implemented before the presentation of the definition serve as a pretest, and learners learn the target items from the corrective feedback. Since trial-and-error practices involve generating a guess, it is associated with the generation effect in psychology, which means generating material tends to lead to better subsequent recall or recognition.

Research on retrieval and trial-and-error practices in second language acquisition

Although the benefits of both types of practices have been well documented in memory research, there is still a need for more research in the realm of second language acquisition. The existing studies on second language acquisition support the conclusion of memory research that retrieval practices help to enhance the memorization of new items. Webb (1921), for example, demonstrated the benefits of retrieval practices in promoting L2 vocabulary learning. All participants first studied Hebrew-English word pairs. Then, half of the participants received a retrieval practice where they were asked to recall the English equivalents of the Hebrew words, and the other half of the participants restudied the word pairs. The results of the similar cued recall test administered after one week showed that performance was better in the retrieval condition than in the restudy condition. Another study conducted by Barcroft (2007) used word-picture pairs as materials. The experimental group was asked to remember pictures with their corresponding L2 words. Then only the pictures were presented, and the learners had to recall the corresponding words. When they finished the practice, the word-picture pairs were displayed again as feedback so that the learners could verify and consolidate the word-picture mapping. The control group, however, studied the same materials twice without the exercise. The result was quite similar to that of Webb’s study, verifying the superior effect of retrieval over simply restudy.

The results regarding trial-and-error practices in promoting L2 knowledge are mixed. Trial-and-error practices were demonstrated to be quite effective in the study of Potts and Shanks (2014). The learning items were either unfamiliar English words or Euskara vocabulary. Participants learned the definitions of these items in three randomly interleaved formats: by generating a response and then being given corrective feedback, by reading the paired associates, or by selecting from two possible choices followed by feedback. The posttest adopted a multiple choice test and the results showed that the condition of generating a response followed by

feedback led to better memory for the correct definitions than the other two conditions. However, the results were challenged because the incorrect options in the posttest were mostly new items, thus generating a guess was more likely to benefit the recognition of form instead of the recall of word-meaning pairings (Strong & Boers, 2019b). Seabrooke et al. (2019) improved the design of the posttest by including the incorrect options that also appeared in the learning materials, and found that generating guesses did not improve the performance. Additionally, in a cued recall test where the meaning was presented and the word was to be recalled, the study-only group even performed better. Thus, the author concluded that trial-and-error practices only strengthen the memory of isolated items rather than paired associates. However, the presentation time was longer in the error. It has been demonstrated in Elgort's study (2017) that participants performed significantly better in a subsequent meaning generation task for novel L2 words when they had correctly inferred the meanings from context than when incorrect or no explicit inferences were made. The effect of incorrect inferences was not worse than that of no inference. Therefore, trial-and-error practices might be more effective in contextual word learning.

Research has only recently begun to compare the effectiveness of retrieval and trial-and-error practices to explore which type of the two practices might be more effective than the other (Strong & Boers, 2019a, 2019b; Elgort, Beliaeva, and Boers, 2020). Committing an error in trial-and-error practices is believed to create confusion which will negatively influence learners' memory. Support could be found in two studies by Strong and Boers where retrieval practices are demonstrated to be superior to trial-and-error ones in enhancing memory of L2 phrasal verbs using gap-fill exercises where particles are missing. However, the study of Elgort, Beliaeva, and Boers on contextual word learning got the opposite result. They found that when compared with the errorless treatment, the trial-and-error treatment resulted in superior declarative knowledge (measured by a meaning recall task) and non-declarative knowledge

(measured by a self-paced reading task) for both L1 and L2 readers. The authors explained that presenting the definition before reading might make learners more familiar with the novel words, thus reducing readers' attention and engagement with contextual cues during reading. Therefore, the answer to which type of practice is more effective might be determined by specific learning situations.

Variables that affect the effectiveness of practices

The effectiveness of practices is affected by multiple variables. For example, practice spacing, provision of feedback, working memory, and modality of input all have an impact on the effectiveness of practice (Lightbown, 2019). Based on previous studies on L2 retrieval and trial-and-error practices, three other important variables are summarized here: practice patterns, item number, and learner proficiency.

Practice patterns

As has been reviewed above, studies on the comparison between retrieval and trial-and-error practices reached different conclusions. One possible explanation for these mixed results might be that different practice patterns were used. Research on the effectiveness of practices in learning L2 multiword expressions has suggested that practices should be errorless to reduce confusion or interference of the initial wrong guesses that are unlikely to be overridden by subsequent corrective feedback. For example, two studies by Boers et al. (2014) and Boers, Dang, and Strong (2016) examined the effects of three formats (word-, letter-, and phrase-formats) of trial-and-error practices on learning L2 verb-noun collocations, with the latter a replication of the former. The results repeatedly verified that learners who worked with the intact phrases had the best performance in memorizing form (Boers et al., 2014; Boers, Dang & Strong, 2016) and meaning (Boers, Dang & Strong, 2016). Similar results were also found in the study of Ferguson et al. (2021). By comparing these three patterns of trial-and-error practices on learning L2 multiword expressions, an assumption could be made that test patterns that lower the initial errors are more beneficial. If this is the case, the errorless retrieval practices are conceivably

more effective in promoting the learning of multiword expressions than the error-prone trial-and-error practices. Empirical support could be found in two studies by Strong and Boers (2019a, 2019b). They directly compared the effectiveness of these two practices in learning L2 phrasal verbs. In the retrieval condition, L2 participants were first presented with the phrasal verb, its English clarification, and an example dialogue. They were then asked to finish the corresponding exercises in the form of completing a dialogue with the particle missing. These two stages were reversed in the trial-and-error condition. Results showed that the retrieval group outperformed the trial-and-error group in both immediate and delayed posttests. In these two studies, L2 learners' responses were almost error-free in the retrieval condition, whereas they were error-prone in the trial-and-error condition.

Similarly, studies on L2 vocabulary learning that used different practice patterns gained different results. A series of studies by Warmington et al. (2013) and Warmington and Hitch (2014) compared errorless and error-prone practices. In these studies, the errorless practices were more like repetition than retrieval: participants were presented with the pictured objects and were told about both the first letter and the entire name, which they were asked to repeat. In the trial-and-error condition, participants were only told about the first letter of the object's name and were asked to guess its entire name. Corrective feedback would be given after they have provided a response. Repetition of the correct names was also required. In the subsequent object naming task, performance of the errorless condition was significantly better in both immediate and delayed posttests. The trial-and-error practice pattern in these studies led to few correct guesses and did not resemble the naturalistic learning situations where complete novel words would be encountered in informative contexts. The studies of Potts and Shanks (2014) and Seabrooke et al. (2019) mentioned above were designed in a similar way where the guesses were made without contextual cues and were almost always incorrect. Studies by Elgort (2017) and Elgort, Beliaeva, and Boers (2020), however, explored the trial-and-error

practices in contextual word learning and attested successfully to the benefit of correct inferences. The inference accuracy of novel words in an informative context was much higher than that of previous studies. The practice of inferring the meaning of unknown words from the context before the presentation of their definitions increases learners' engagement with contextual cues and deepens their encoding.

In sum, when target items are not embedded in an informative context, the errorless practice patterns are generally more beneficial than the error-prone ones, especially in learning novel multiword expressions that require associations not only between form and meaning but also between the constituent items within the multiword expressions (Strong & Boers, 2019b). In addition, the learning of multiword expressions is more difficult for L2 learners because they tend to seek meaning in individual words rather than in chunks (Ferguson et al., 2021). The error-less practice patterns such as presenting a list of intact choices to choose from are more beneficial because they help to reduce confusion and wrong associations that occur in initial guesses, hence enhancing the learning. However, when learners are provided with contextual information and are asked to infer novel words' meaning, the more difficult trial-and-error practices seem to be more favorable than errorless retrieval practices.

Item number

The number of items tackled in one set of practices is one of the important factors that affect learning gains. The factor influences the cognitive burden of the learning process as well as the timing of corrective feedback for trial-and-error practices. Schmidt and Bjork (1992) held that, when compared with immediate (continuous) feedback, delayed (summarized) feedback might slow down initial learning but yield better retention in the long term. Strong and Boers (2019b) explored the effect of item number per practice in both retrieval and trial-and-error practices. In one condition, the feedback followed immediately after each response, and in the other condition, the feedback followed when a whole practice of 14 items was completed. Learning

one item at a time tends to be effortless, while learning several items in a set is more challenging and effortful. Following the view that it is the cognitive effort invested in practices that helps to entrench knowledge, the researchers hypothesized that the practice containing 14 items in a set required more cognitive effort, hence leading to more long-term retention. However, their finding was inconsistent with this hypothesis. In both the immediate and delayed posttests, there was no significant effect for the number of items or the interaction between learning methods (retrieval and trial-and-error) and item number. Note that a floor effect appeared in the delayed posttest, which was not observed in their prior study, where items were learned seven in a row and the retrieval group significantly outperformed the trial-and-error group. Taken the two studies together, it is assumed that 14 items per set brings an overly heavy learning burden, while one-at-a-time invites insufficient effort to reap the rewards of retrieval, and a practice of 7 items might be optimal by inviting the desirable difficulty for those participants. Since no conclusive picture has yet emerged about the effect of item number, more replications are needed in this respect.

Learner proficiency

Individual differences such as working memory and proficiency play an important role in affecting the effectiveness of practices. The studies by Elgort (2017) and Elgort, Beliaeva and Boers (2019) explored the role of learner proficiency in contextual word learning. Elgort (2017) found that the detrimental effect of incorrect inferences on subsequent meaning recall was more significant for less proficient L2 learners. Similarly, Elgort, Beliaeva and Boers (2019) recruited native English speakers and Chinese English learners and found that knowledge gains of the pseudo-words were greater for the L1 participants than the L2 participants. Also, the effect of attrition over time was greater for L2 participants, and only L1 participants showed an advantage for incorrectly inferred items in the trial-and-error treatment. These findings support the conclusion that high-proficiency learners establish more robust lexical-semantic representations than less proficient learners. Therefore, at least in the circumstances of these two studies, errorless

retrieval practices may be more appropriate for low-proficiency learners, while error-prone trial-and-error practices are more beneficial for high-proficiency learners. In these studies, learner proficiency is related to the ability to correctly infer meaning from the contextual information. As proposed by Krashen (1982, p63), "the optimal input is comprehensible." Compared with the less proficient learners, learners of higher proficiency are more likely to comprehend the input and correctly infer the meaning of the novel words and, ultimately, benefit more from initial inferences.

General Discussion

This review provides a brief overview of the recent literature concerning studies on retrieval and trial-and-error practices in second language acquisition. Both retrieval and trial-and-error practices have been demonstrated as more effective in promoting memorization of novel items than simply restudying. Although retrieval practices lead to better performance than trial-and-error practices in most cases, a conclusion could not be drawn safely because the effectiveness of practices largely depends on specific learning situations. The effectiveness of practices in promoting L2 learning is affected by multiple variables. However, the interactive effect between these variables is still underexplored.

The findings of previous studies suggest that practices that are not too difficult or too easy lead to the best learning. If the practices are easy to answer, learners will probably lose interest and not attend, while highly difficult practices will result in more errors that will interfere with the subsequent learning. The findings can be systematically explained from the perspective of Bjork's desirable difficulty framework (2018) in cognitive psychology. It proposes that procedures that pose certain difficulties and slow the speed of learning will lead to longer retention. Based on the desirable difficulty framework and the L2 cognitive difficulty framework, Suzuki, Nakata, and Dekeyser (2019, 2020) presented a unified framework for researching L2 practices. Their framework provides theoretical support for research on retrieval and trial-and-error practices by illustrating how practice

conditions, linguistic difficulty, and individual differences interact to influence desirable difficulty during practices. In the studies on retrieval and trial-and-error practices reviewed above, the primary difference, that is, the placement of the definition, and other variables, including the practice patterns and number of items per practice, belong to the practice condition. Different target items that are different in degrees of difficulty, such as vocabulary and multiword expressions, are associated with linguistic difficulty, and learner proficiency is included in individual differences. These multiple variables interact and determine learners' perceived difficulty towards a practice and, consequently, the effectiveness of practices. Effortless practices can lower initial errors and interference during learning but invite insufficient cognitive effort that enhances long-term memorization. Effortful practices require more cognitive effort to get over the interference caused by initial incorrect guesses or to infer meaning from contexts. The learning will be entrenched in the long run if learners are able to handle the challenging practice with effort, but when the learning burden is too heavy, the practice will inhibit learning. Therefore, a specific type of practice that helps one group of learners enhance memory might inhibit the learning of another group of learners at a different proficiency level, or the same practice targeted at different learning materials might affect the learning differently. The mixed results of studies on L2 practices with different participants could be well explained under the desirable difficulty framework. Practice conditions should be designed to match the difficulty of learning materials and learner internal factors to induce the optimal difficulty. Less demanding practice conditions are appropriate for learning more complex targets. For example, when learning multiword expressions that are quite difficult for L2 learners, it is recommended to present the definition before practice, keep the multiword expressions intact during practices, and provide feedback after having learned an appropriate number of items. However, more demanding practice conditions are ideal for learning less complex targets. In contextual word learning, the more effortful trial-and-error practices may be

better at helping learners engage in reading, establish memory traces, and ultimately promote form-meaning connections.

Implication and Limitations

The research on the practices applied in second language acquisition has implications for textbook design and practice implementation in the classroom. Teachers are recommended to implement appropriate practices based on the difficulty of learning targets and students' individual differences by controlling practice conditions such as practice patterns, feedback timing, and task complexity to invite the optimal effort. Teachers can also change the procedures of exercises that are presented in a textbook for different learners. However, since there is no objective measure of learning difficulty and multiple variables are intertwined, implementing the optimal practice that induces the desirable difficulty remains a challenge for education.

The existing studies are not without limitations. One limitation is that studies on learning L2 novel items through practice are still quite limited in number, and no clear conclusion can be reached at present. Practice conditions, linguistic difficulty, and individual differences all affect the learning gains through practice, and only a few studies have investigated the interaction between them. More replications are urgently needed to find out how these factors work together. The second limitation is that activities implemented in most of these studies are different from those used in authentic learning situations, where more activities are included in learning the same target item. Future studies will be welcomed where additional activities are included to ensure that the activities resemble those used in an authentic classroom context. The third limitation is the lack of various measurement methods. More fine-tuned measurement is needed to measure different kinds of knowledge, such as explicit and implicit knowledge. For example, ERP is more sensitive to learners' improvement than the test patterns used in these studies, and may shed some light on the effect of different practices on different cognitive processes.

References

- [1] Bjork, R. A. "Being suspicious of the sense of ease and undeterred by the sense of difficulty: Looking back at Schmidt and Bjork (1992)." *Perspectives on Psychological Science* 13 (2018): 146–148.
- [2] Boers, F., Dang, T.C.T., & Strong, B. "Comparing the effectiveness of phrase-focused exercises: A partial replication of Boers, Demecheleer, Coxhead, and Webb (2014)." *Language Teaching Research* 21 (2017): 362–380.
- [3] Boers, F., Demecheleer, M., Coxhead, A., & Webb, S. "Gauging the effects of exercise on verb–noun collocations." *Language Teaching Research* 18 (2014): 54–74.
- [4] DeKeyser, R. M.. *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. New York: Cambridge University Press, 2007.
- [5] Elgort, I. "Incorrect inferences and contextual word learning in English as a second language." *Journal of the European Second Language Association* 1(2017): 1–11.
- [6] Elgort, I., Beliaeva, N., & Boers, F. "Contextual word learning in the first and second language: definition placement and inference error effects on declarative and nondeclarative knowledge." *Studies in Second Language Acquisition* 42(1) (2020): 7-32.
- [7] Ferguson, P., Siyanova-Chanturia, A., & Leeming, P. "Impact of exercise format and repetition on learning verb–noun collocations." *Language Teaching Research*, 0(0); (2021). <https://doi.org/10.1177/13621688211038091>
- [8] Krashen, Stephen. *Principles and Practice in Second Language Acquisition*. Pergamon Press, 1982.
- [9] L. W. Webb. "A Comparison of Two Methods of Studying with Application to Foreign Language." *The School Review* 29:1 (1921): 58-67
- [10] Lightbown, P. "Perfecting practice." *Modern Language Journal* 103(2019): 703–712.
- [11] Roediger, H. L., III, & Karpicke, J. "The power of testing memory: Basic research and implications for educational practice." *Perspectives on Psychological Science* 1(2006): 181–210. <https://doi.org/10.1111/j.1745-6916.2006.00012.x>
- [12] Schmidt, R. A., & Bjork, R. A. "New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training." *Psychological Science* 3(1992): 207–217.
- [13] Seabrooke, T., Hollins, T. J., Kent, C., Wills, A. J., & Mitchell, C. J. "Learning from failure: Errorful generation improves memory for items, not associations." *Journal of Memory and Language* 104 (2019): 70–82. <https://doi.org/10.1016/j.jml.2018.10.001>
- [14] Strong, B. and Boers, F. "The Error in Trial and Error: Exercises on Phrasal Verbs." *TESOL Q* 53(2019a): 289-319. <https://doi.org/10.1002/tesq.478>
- [15] Strong, B. and Boers, F. "Weighing Up Exercises on Phrasal Verbs: Retrieval Versus Trial-and-Error Practices." *The Modern Language Journal* 103(2019b): 562-579. <https://doi.org/10.1111/modl.12579>
- [16] Suzuki, Y., Nakata, T. and Dekeyser, R. "Empirical Feasibility of the Desirable Difficulty Framework: Toward More Systematic Research on L2 Practice for Broader Pedagogical Implications." *The Modern Language Journal* 104(2020): 313-319. <https://doi.org/10.1111/modl.12625>
- [17] Suzuki, Y., Nakata, T., & DeKeyser, R. M. "The desirable difficulty framework as a theoretical foundation for optimizing and researching second language practice." *Modern Language Journal* 103(2019): 713–720.
- [18] Warmington, M., & Hitch, G. J. "Enhancing the learning of new words using an errorless

learning procedure: Evidence from typical adults." *Memory* 22(2014): 582–594.
<https://doi.org/10.1080/09658211.2013.807841>

- [19] Warmington, M., Hitch, G. J., & Gathercole, S. E. "Improving word learning in children using an errorless technique." *Journal of Experimental Child Psychology* 114(2013): 456–465.
<https://doi.org/10.1016/j.jecp.2012.10.007>