Is perceiving another’s error detrimental to learning from corrective feedback?

Nobuyoshi Iwaki1,*, Mizuki Tomisawa1, Reika Suzumori1, Akira Kikuchi1, Isao Takahashi2, Saeko Tanaka3 and Susumu Yamamoto1

Abstract: Students frequently make errors when learning, but generating errors can enhance subsequent learning from corrective feedback (i.e. the testing effect). There have also been cases in which errors generated by someone can be shared with others. Notably, it has been found that listening to errors provided by someone could hinder the other’s subsequent learning from corrective feedback. In this study, we investigated whether the learning of someone who knows the correct answers beforehand is impaired by perceiving another’s error (i.e. retroactive interference). Thus, two experiments were conducted: a simulated question-and-answer situation (Experiment 1) and a mutual question-and-answer situation in which 2 students quizzed one another (Experiment 2). The results showed that the cued-recall performance of the tester who had received incorrect answers a week prior was not different from that in the case of learning with only reading (control condition), showing no negative effects. Furthermore, the memory performance of participants who took a test was better than the control condition, demonstrating the testing effect. We, thereby, discuss the efficacy of the reciprocal testing procedure as a learning method.

Subjects: Teaching Methodology & Practice; Psychological Science; Educational Psychology

ABOUT THE AUTHORS
Our research team is interested in practical and theoretical issues concerning the testing effect/retrieval-based learning and learning from errors. Specifically, we have investigated several procedures to enhance the effect of corrective feedback on learning. Recently, we have focused on the relationship between memory for errors and error correction.

PUBLIC INTEREST STATEMENT
It has been repeatedly confirmed that retrieval via testing has a robust ability to enhance learning (resistance to forgetting). A body of studies has attempted to apply retrieval-based learning to educational settings, and, at the same time, they have been successful. However, students cannot enjoy individual testing due to fatigue, and then, consequently, learning may cease. Mutual testing with peers (i.e. questions and answers) may motivate taking the opportunity to enjoy the testing procedure. However, mutual testing with peers has been a concern. It is possible that when a tester receives erroneous responses from a respondent, the tester’s memory for correct answers might be hindered. This possibility was investigated in the present study, which includes two experiments, and neither experiment confirmed such a negative effect. Thus, mutual testing with peers seems a suitable procedure for students to enjoy the benefit of testing.

© 2020 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.
1. Introduction

Some errors occur when people learn new materials or skills. Learning performance usually begins with the occurrence of many errors and materials and skills are gradually acquired. Additionally, in educational settings, erroneous information generated by a certain learner is often shared by other learners. In other words, a learner functions as a source for supplying false information. In the following, we sort findings in the literature based on (a) whether false information is generated by learners for themselves or supplied from an external source and (b) the experienced order of correct and incorrect information.

According to the traditional behaviourism theory, generating errors for oneself could impair and delay the acquisition of correct responses. For example, when students learn words and their definitions, it is likely that a learner recalls a false word from the definition as the clue or uses incorrect spelling in the early process of learning. Skinner (1958) claimed that in the early phase of learning, many hints are helpful to induce correct responses and prevent learners from making errors. This claim is based on the idea that generating an incorrect response to a clue strengthens the association and hinders making new associations between the clue and the correct response. However, many empirical findings do not agree with Skinner’s claim.

If a person recalls previously learned materials, that information is more likely to be remembered later (Karpicke & Roediger, 2008; Roediger & Karpicke, 2006a). This phenomenon is termed the “testing effect” or “retrieval-based learning” (Delaney, Verkoeijen, & Spirgel, 2010; Karpicke, Lehman, & Aue, 2014; Roediger & Karpicke, 2006b; Roediger, Putnam, & Smith, 2011). Importantly, even if a question receives an incorrect answer, it can still potentiate the learning of information provided through corrective feedback, and this has been well established (Huelser & Metcalfe, 2012; Izawa, 1970; Kornell, Hays, & Bjork, 2009; Kornell & Vaughn, 2016; McDermott, Agarwal, D’Antonio, Roediger, & McDaniel, 2014; Richland, Kornell, & Kao, 2009; Slamecka & Fevreiski, 1983). In brief, retrieving an incorrect answer could enhance the subsequent learning of corrective feedback information. However, this phenomenon seems to be limited to cases in which people generate an answer by themselves. False information may be provided by an external source, and, in this case, studies have suggested the possibility that subsequent learning is adversely affected.

Grimaldi and Karpicke (2012) conducted paired-associate learning experiments using weakly associated word pairs. They established a pre-test condition in which participants guessed a potential target word from a cue word and a constrained-pre-test condition in which they completed word stems when provided fragments of a lure word. These activities were followed by the provision of corrective feedback (i.e. correct word pair). In the no-pre-test control condition, participants experienced only correct word pairs (Experiment 2). As the forward association from a cue to target was 5%, most responses were erroneous, even in the pre-test condition, and these error trials were analysed. The results demonstrated that the proportion of correct answers in the final cued-recall test was better for the pre-test condition than for the control condition, and the proportion of correct answers for the constrained-pre-test condition, in which the participants were forced to retrieve lures, was worse than that for the control condition. As Metcalfe and Xu (2018) suggested, it could be thought that the lures of the constrained-pre-test condition were provided by the experimenter; therefore, it was suggested that the perception of false information from an external source negatively affected the subsequent learning of the corrective feedback information.

As mentioned, learners are considered an external source for supplying false information as well. In Metcalfe and Xu’s (2018) study, the experimenter provided three participants with general
knowledge quizzes online. A participant appointed as the respondent answered questions (call-on /self condition). The other two participants heard the respondent’s answer (call-on/other condition). The experimenter subsequently furnished the correct answer. Additionally, in the control condition, the experimenter read the quiz and its correct answer aloud, and the participants merely listened to them (listening condition). Notably, the proportion of correct answers in the final cued-recall test in Experiment 1 was lower for the condition in which the respondent’s answer was heard (i.e. call-on/other) when compared with the control condition. Therefore, other’s wrong answers could hinder the perceiver’s learning.

The two aforementioned experiments are cases in which correct answer feedback was given to participants after they had received false information that originated from an external source (experimenter and other respondents). Thus, this negative effect could be understood based on the proactive interference of the classical interference theory (Underwood, 1957; Underwood & Freund, 1968). Schematically, if a cue (A) was previously associated with a respondent’s error (B), the association between (A) and (B) might interfere with the association between the cue (A) and the correct answer (C).

Notably, it is also possible that the individual who knows the correct answers beforehand experiences false information later. Based on the classic interference theory, the association between a cue (A) and target (B) could be hampered by a subsequent experience (C), called “retroactive interference” (Briggs, 1957; Postman & Underwood, 1973). However, in reality, it is difficult to imagine a case in which a person, having known the correct answers, generates false information purposely by themselves. A related study was reported by Lehman and Karpic (2016). In their Experiments 4 and 5, during the first phase, participants learned target words. In the second phase, participants were given a cue word, which was semantically related to a target word, along with stems of several cue-related words, including the target. The number of cue-related words, except for the target (by their term, “mediator”), was the independent variable. The results of both experiments demonstrated that the cued-recall performance was the best when the least amount of false information (one mediator) was generated (in Experiment 5, the more the false information was generated, the more the memory performance decreased monotonically). These findings suggest that if the learning of targets was followed by the generation of cue-related information that differed from the target, the target became subsequently less likely to be recalled from a cue. However, these experiments did not have a condition in which no false information (lure) was generated, and participants had not been informed that the words (targets) presented for learning in the first phase and the cue words given in the second phase constituted the true (correct) word pair. Therefore, it is unknown whether a memory of a previously acquired correct word pair could be impaired by the subsequent generation of false information. As mentioned, however, it seems unlikely that a person who knew correct answers in advance would intentionally generate false information.

Contrastingly, it is common that a person having correct information beforehand receives false information from an external source. A scene in which a student provides a wrong answer to a question posed by a teacher is typical. Students who know the correct answer should be present in the classroom; therefore, it is theoretically possible that a memory for a correct answer could be retroactively interfered with by false information subsequently provided by other students. Another case is given by “peer assessment”, an arrangement in which individuals consider the amount, level, value, worth, quality or success of the products or outcomes of learning of peers (Topping, 1998). Students judge a product made by the other peers quantitatively (e.g. scores) and/or qualitatively (e.g. comments) based on previously provided criteria, and they present the peer with a feedback (Falchikov, 1995; Gielen, Peeters, Dochy, Ongena, & Struyven, 2010; Hovardas, Tsivitanidou, & Zacharia, 2014). Importantly, because peer assessment is essentially a formative assessment conducted with the feedback among students, false information included in the content of the one assessed is shared among peers. The assessor’s memory for correct answers might be impaired by experiencing false information during the review of the content. However, according to our review of the literature, it is unknown whether such a negative effect actually works.
It is also possible that encoding other’s errors enhances learning. Specifically, when an error given from others is encoded as a mediator linking a cue and a target, and the error can be used as an additional retrieval cue, cued-recall performance for correct answers later might be improved (Barnes & Underwood, 1959; Carpenter, 2011; Pyc & Rawson, 2009). However, it is unknown whether the effect might occur when a person who has known correct answers in advance receives erroneous responses from the other. Our goal is the investigation of this issue.

The present study examines whether the learning of a person who already knew the correct answers might be impaired by perceiving false information from an external source. In two experiments, we simulated question-and-answer situations because when participants play the role of the tester, they know the correct answers before asking a respondent. In Experiment 1, we used a situation in which a participant presented a quiz to another fictitious participant and received an answer from them. In Experiment 2, two people performed questions and answers reciprocally. The important argument was whether the learning of a tester given false information from an external source would be impaired when compared with a control condition (read-only condition and condition in which a tester is given correct information). We also examined another prediction. In the fill-in-the-blank condition, participants played the role of a respondent during testing and took notes. In this condition, the retrieval practise should enhance learning, resulting in improved memory when compared with the case in which they learned materials through reading only (the testing effect).

2. Experiment 1

2.1. Method

2.1.1. Participants
From Iwate University, 24 undergraduate and graduate students (8 males and 16 females, mean age = 20.4, SD = 1.1) participated in the experiment for monetary compensation. The participants were treated according to the ethical guidelines of the American Psychological Association and provided written informed consent.

2.1.2. Materials
The experimenters collected 48 short statements from general psychology textbooks. These materials were created from various psychology fields (see the Appendix A). The mean length of the statements was 73 characters (range: 64–83) in Japanese notation (examples are shown in the Appendix A). Of the 48 statements, 12 were assigned to the read-only condition, 12 to the fill-in-the-blank condition, and the remaining 24 to the tester condition, randomly. Regarding the tester condition, correct answers were given to the tester in half of the 24 trials (tester/correct answer condition), whereas wrong answers were given (tester/wrong answer condition) in the remaining trials. These conditions, therefore, operated as a within-participant variable. Two wrong answers for each quiz were collected in a pilot study and randomly selected for each participant.

2.1.3. Task and procedure
During the initial learning session, a participant was provided with three pieces of paper printed with 48 statements. The participant was encouraged to study the materials at their own pace after they received a retest notice. Two checkboxes were set up at the end of each statement. The total time required to complete the readings was measured. Then, after a 2-minute distraction task in which they wrote down the names of countries, the participants were ready to enter the experimental treatments.

A summary of the experimental procedure is shown in Figure 1. A participant sat in front of a monitor, and the experimenter operated a computer that faced away from the participant. The material was presented in white on the monitor, with a black background. During the read-only trials, a statement with the target word in parentheses was displayed. The participant was
instructed to read the text aloud and press the Esc key to show the target word displayed in yellow below the text for 5 seconds and read it aloud.

In the fill-in-the-blank condition, while the statement was being shown, the target word was replaced with only the first letter of the word, followed by an appropriate number of blank spaces. The participant then read the text aloud and took notes on the target word. When the participants could not remember the target, they inserted a question mark ("?"). There was no time limit for responding. Following the participants’ response, they pressed the Esc key to show the correct answer in yellow below the text for 5 seconds, and then they were instructed to read it aloud.

In the tester conditions, participants were instructed to repeat the statement on the monitor and relay it through a microphone to another participant waiting in the next room (the other participant was not actually present). The participants assigned as the tester were asked to read the target silently. The experimenter pushed the Esc key at the same time as the participant finished reading the text. As soon as the experimenter pressed the key, a correct or incorrect answer determined by the computer was displayed under the text for 5 seconds after an inconsistent time lag. The duration of this delay was randomly altered between 4–9 seconds across trials to vary the timing of the presentation of responses. It was emphasized that correct answers were shown in yellow, and wrong answers were shown in white. Participants read this aloud. Following this, in the tester/correct answer trials, the phrase “That is correct” was presented for 2 seconds on the monitor, and participants were instructed to read it aloud. In the tester/wrong answer trials, a correct answer was displayed for 5 seconds, and the participants repeated it to the other participant who was actually absent. As mentioned, reading both the text and the correct
answer feedback aloud was common in all conditions. In the tester/wrong answer condition, the participants read the wrong answers as well.

The order of the materials was randomised for each participant in both the initial learning session and the experimental treatments. The order of read-only, fill-in-the-blank and tester (with either a correct or an incorrect answer) conditions was counterbalanced across the participants. Prior to a trial block for each condition, the participants received two practise trials with new materials. One week later, the final test was administered using three pieces of paper printed with 48 statements, and the parts of target words were kept blank. In the final test, participants responded with an answer sheet that had no hints for the target words. The order of the materials was randomised again for each participant. At the end of the experiment, participants were debriefed and thanked for their participation.

2.1.4. Statistical analysis
Three comparison analyses were conducted. First, to confirm the testing effect, the fill-in-the-blank condition was compared with the read-only condition. Second, to examine the effect of perceiving the other’s error on the tester’s learning, the tester/wrong answer condition was compared with both the read-only and tester/correct answer conditions. The experiment used the Bonferroni method for multiple comparisons and set \( \alpha = 0.0167 \) (= 0.05/3). Hedges’ \( g \) was used as the measure of effect size for all t-tests.

2.2. Results and discussion
The time required to perform the initial learning session was \( M = 30.2 \) minutes (\( SD = 11.2 \)). The proportion of correct responses in the initial test for the fill-in-the-blank condition was \( M = 0.61 \) (\( SD = 0.20 \)). The proportion of intrusion errors in the final test that was the same as those in the initial test for the fill-in-the-blank condition was \( M = 0.03 \) (\( SE = 0.01 \)), and that of incorrect answers given by the other person (i.e. experimenter) for the tester/wrong answer condition was \( M = 0.02 \) (\( SE = 0.01 \)).

Table 1 shows the mean proportions of correct responses for each condition in the final test. The proportion for the fill-in-the-blank condition was the highest (\( M = 0.48 \), \( SE = 0.05 \)), and those for the other three conditions were similar to one another (read-only, \( M = 0.33 \), \( SE = 0.04 \); tester/correct answer, \( M = 0.30 \), \( SE = 0.03 \); tester/wrong answer, \( M = 0.32 \), \( SE = 0.05 \)). In multiple comparison tests, the proportion for the fill-in-the-blank condition was significantly higher than that for the read-only condition (\( t (23) = 3.95, p < 0.001, g = 0.66 \)). The proportion for the tester/wrong answer condition did not differ significantly from those for either the read-only condition (\( t (23) = 0.15, g = 0.03 \)) or the tester/correct answer condition (\( t (23) = 0.64, g = 0.13 \)).

The proportion of correct answers in the final test was significantly higher for the fill-in-the-blank condition when compared with that for the read-only condition. This result confirms the testing effect. Furthermore, the correct proportion for the tester/wrong answer condition was similar to that for the read-only (\( g = 0.03 \)) and the tester/correct answer (\( g = 0.13 \)) conditions. As the effect sizes were minute, no substantial difference was observed between these conditions. These results did not support the idea that the learning of a tester who knew the correct answer is retroactively interfered with by the other’s error (i.e. false information from an external source). In addition, these results are inconsistent.

### Table 1. Mean proportions of correct responses for each condition in the final test for Experiments 1 and 2. Standard errors are in parentheses

|                | Read-only | Fill in the blank | Tester/CA | Tester/WA |
|----------------|-----------|-------------------|-----------|-----------|
| **Experiment 1** |           |                   |           |           |
| Read-only      | 0.33 (0.04) | 0.48 (0.05)       | 0.30 (0.03) | 0.32 (0.05) |
| **Experiment 2** |           |                   |           |           |
| Read-only      | 0.44 (0.05) | 0.58 (0.04)       | 0.45 (0.05) | 0.46 (0.05) |

CA: correct answer, WA: wrong answer
with the semantic mediation hypothesis. If the information of errors might work as a mediator between a cue and a target, then a positive effect should have been confirmed. When errors were not self-produced but given from an external source, these errors might not work as a mediator.

However, this experiment has a limitation. Regarding tester conditions, participants read the statement aloud to the other (fictitious) person, but in the fill-in-the-blank condition, they were shown the text on the monitor without any sound. The presentation of materials was thus not the same for each condition. In a typical procedure during questions and answers, however, the roles of two learners are replaced between the fill-in-the-blank condition and the tester condition; both the quiz presentation and their answers may often be performed orally. Considering this point, we modified the procedure in Experiment 2 to reconfirm the results of Experiment 1.

3. Experiment 2

3.1. Method

3.1.1. Participants
A total of 30 new students from Iwate University (9 males and 21 females, mean age = 21.2, SD = 1.8) were recruited. They were paid for their participation and provided written informed consent. These participants constituted one member of a pair during the questions and answers. The experiment used a confederate as the other member of the pair, but this information was not provided to participants.

3.1.2. Materials, design and statistical analysis
Most of the same materials used in Experiment 1 were used again in Experiment 2. However, three materials were replaced in Experiment 2 because the proportions of answers for these materials were found to be incorrect in the initial and final tests. Of the 48 statements, 16 were assigned to the read-only condition, 16 to the fill-in-the-blank condition and 16 to the tester condition, with either a correct or an incorrect answer. In the tester condition, one-half of the materials were assigned to the correct answer trials, and the other half was assigned to the wrong answer trials. We pooled the wrong answers used in Experiment 1 and the newly occurring errors, and three wrong answers were selected for each material before the experiment. The design and statistical analysis of the experiment were the same as in Experiment 1.

3.1.3. Task and procedure
Differences in the procedure from Experiment 1 are as follows: The mean time required to complete the initial learning session in Experiment 1 was approximately 30 minutes, and 20 of the 24 participants (83.4%) finished the initial session within 40 minutes. This experiment required participants to study all materials for 40 minutes and read them at least twice. Following the initial learning session, the participants performed a 2-minute distraction task.

The summary of the procedure is shown in Figure 2. The experimenter controlled the presentation of the materials on the monitor. In the read-only condition, participants read the text aloud and then pressed the Esc key. When the monitor displayed the target word in yellow for 5 seconds, the participants were instructed to read it silently (in Experiment 1, the participants read the word aloud). During this procedure, the confederate waited in the next room.

Questions and answers were administered for the fill-in-the-blank and the tester condition. Participants and the confederate sat near one another but were separated by a partition so that they could not see each other’s monitors. However, they could hear each other’s voices clearly. In the fill-in-the-blank condition, the confederate who was assigned as a tester read the statement displayed on the monitor aloud to the participant but read the target word silently. The text was not visible on the participants’ monitor. Although the initial character of the target word was also presented on the monitor in Experiment 1, this was not done in Experiment 2. After the participants wrote their response
on an answer sheet, they reported their answer orally to the tester (confederate). If their answer was correct, the tester pressed the Esc key to display the correct answer in yellow on both of their monitors and told the participant, “That is correct.” In the case of an incorrect answer, the tester responded by telling the participant the correct answer.

The procedure of the tester conditions with a correct or incorrect answer was the same as that of the fill-in-the-blank condition, in which their roles were replaced (participants played the role of the tester). The confederate assigned as a respondent noted an answer that was ordered beforehand by the computer and reported it orally. Concerning incorrect answers, one of three pooled words for each material was selected randomly by the computer for each participant.

### 3.2. Results and discussion

The proportion of correct answers in the initial test for the fill-in-the-blank condition was $M = 0.67$ ($SD = 0.25$). In terms of intrusion errors in the final test, the proportion of participants’ self-generated errors for the fill-in-the-blank condition was $M = 0.01$ ($SE = 0.003$), and the proportion of incorrect answers provided by the confederate assigned as the respondent for the tester/wrong answer condition was $M = 0.004$ ($SE = 0.004$).

Mean proportions of correct responses for each condition in the final test are shown in Table 1. The fill-in-the-blank condition had the best score ($M = 0.58$, $SE = 0.04$). The other three conditions had similar results to one another (read-only, $M = 0.44$, $SE = 0.05$; tester/correct answer, $M = 0.45$, $SE = 0.05$; tester/wrong answer, $M = 0.46$, $SE = 0.05$). Multiple comparison tests demonstrated that the proportion of correct answers for the fill-in-the-blank condition was significantly better than that for the read-only condition ($t(29) = 6.27$, $p < 0.001$, $g = 0.57$). The proportion for the tester/
wrong answer condition was not significantly different from those for the read-only condition \((t_{(29)} = 0.98, g = 0.09)\) or the tester/correct answer condition \((t_{(29)} = 0.32, g = 0.05)\).

In Experiment 2, a participant and the confederate communicated with one another orally, and their roles were exchanged between the fill-in-the-blank and the tester conditions. The procedure was symmetrical between these conditions, and the experimental situation for questions and answers was more realistic than that in Experiment 1. Moreover, the results of Experiment 2 reproduced those of Experiment 1.

4. General discussion

This study investigated whether the learning of a person who knew the correct answers beforehand might be impaired by perceiving false information provided by an external source. Two experiments provided a similar pattern of results, demonstrating that the testing effect occurred even in the questions and answers and that the tester’s learning was not hindered or facilitated by receiving false information originating from an external source.

Metcalfe and Xu \((2018)\) suggested that perceiving other’s generated errors would reduce the perceiver’s learning outcomes when compared with the case of the control condition. Notably, if a learner heard another’s wrong answer before receiving corrective feedback in a situation, the cued-recall performance was later lowered when compared with the case of the control condition, in which the participant merely heard the question and answer. In educational settings, such a situation is similar to a case in which a student incorrectly answers a question posed by a teacher, and the other students hear the wrong answer without thinking of an answer for themselves. In a question-and-answer situation, a tester knows the correct answer before quizzing a respondent. In this case, either a negative effect (i.e. retroactive interference) predicted by a classical interference theory \((Briggs, 1957; Postman & Underwood, 1973)\) or a positive effect via the semantic mediation of errors between a cue and a target \((Barnes & Underwood, 1959; Carpenter, 2011; Pyc & Rawson, 2009)\) might occur. Our experiments showed neither a negative nor a positive effect. In the tester/wrong answer condition, false information prepared in advance was displayed on a monitor in Experiment 1, and it was relayed to the participants by the confederate in Experiment 2. There did not seem to be any effect from false information given externally, regardless of modality.

The findings can be sorted as follows: (a) when a learner generates a wrong answer from a cue followed by receiving corrective feedback, the learning of the feedback information is enhanced \((e.g. Kornell et al., 2009; Richland et al., 2009)\); (b) if a learner perceives false information from an external source followed by corrective feedback, the learning may be impaired \((e.g. Grimaldi & Karpicke, 2012; Metcalfe & Xu, 2018)\); (c) in a case in which a learner knew the correct answer before receiving false information externally, their memory for the correct answer appears not to be impaired or enhanced \((the\ finding\ of\ this\ study)\); (d) when a learner knew the correct answer before generating false information for themselves, how the false information influences their memory for the correct answer is unclear \((however,\ this\ case\ seems\ to\ be\ rare\ in\ reality)\).

4.1. Educational implications and future directions

A reliable testing effect was observed in Experiment 2, in which two persons performed a question-and-answer procedure among themselves. The question-and-answer procedure is thus considered valuable as a learning method, and it seems likely that both learners can benefit from reciprocal testing without hindrances. Another benefit of questions and answers test is their low cost, i.e. learners can use learning materials they already have, and teachers can save labour and time. Furthermore, learners can practise questions and answers reciprocally anytime, even during class. Additionally, questions and answers are low tech. Although it has been demonstrated that applying a computer-based programme to retrieval-based learning is factually effective \((Grimaldi & Karpicke, 2014)\), learners must move to a room with a computer, and teachers must prepare materials for the programme beforehand. Questions and answers require only a peer and study materials \((e.g. a\ textbook)\). The procedure can be conducted flexibly anywhere, particularly at school, and is particularly suited to reinforcing computer-based lessons.
In addition, the evidence that learners respond overtly during the questions and answers is also important. A study has demonstrated that the magnitude of the testing effect can be modulated through response formats (Tauber et al., 2018). The results of such investigations emphasise that for the testing effect to occur, participants must perform exhaustive retrieval. If the participants respond overtly (e.g. oral reporting, note-taking), they actually execute a comprehensive retrieval practise; however, if they respond covertly (i.e. mental recall), the retrieval practise may cease along the way or may not be executed at all in some instances. Studies using paired-associate learning have demonstrated that the difference in response formats (overt vs. covert) did not modulate the testing effect (Putnam & Roediger, 2013). Moreover, if any moderation was found, the influence was minimal (Jönsson, Kubik, Sundqvist, Todorov, & Jonsson, 2014). Conversely, when dense, information-heavy learning materials, such as key term definitions, were used, as demonstrated in the study conducted by Tauber et al. (2018), covert responses did not lead to a significant testing effect. Researchers thus recommend an overt response while learning information-rich material. Additionally, no such concern exists regarding questions and answers because the procedure requires people to respond overtly (e.g. oral reporting).

A limitation of this study is that both experiments asked participants to provide their answers in writing. In the usual question-and-answer situation, learners report their answer orally. The procedure is simple because only the learning materials are required. A concern to be investigated is the magnitude of the testing effect when making participants answer orally without writing. Presumably, even in such a case, our findings would be reproduced. Sundqvist, Mäntylä, and Jönsson (2017) found that the magnitude of the testing effect was similar among the response formats of writing or typing. Putnam and Roediger (2013) found almost the same magnitude of the testing effect among participants responding through typing or an oral report. In addition, the present study took a within-participant approach to examine the conditions employed. To avoid any carry-over effect, experiments using a between-participants design should be conducted in the future research.

Finally, Metcalfe and Xu (2018) suggested that if learners are provided with the opportunity to think about their answer before perceiving someone’s erroneous answer, a negative effect (i.e. proactive interference) could be prevented. This suggestion is critical when attempting to apply the question-and-answer procedure to a situation with three or more learners. When a tester poses a question to two or more learners, one learner may hear the other learner’s erroneous response before receiving a correct answer. Furthermore, a worthwhile investigation is the efficacy of using the combination of elaborate study and questions and answers in teaching. Useful information would be whether they have the same level of effect as an elaborative strategy, followed by individual testing. Moreover, the applicability of the results to participants of various ages should be examined.

Funding
This work was supported by JSPS KAKENHI grants (16K04285, 16K04331) and Iwate University Grant for High Level Research (2017).

Author details
Nobuyoshi Iwaki 1
E-mail: iwaki@iwate-u.ac.jp
Mizuki Tomsawa 1
E-mail: signal2810@gmail.com
Reika Suzumori 3
E-mail: suzumoryu219529@oazio.jp
Akira Kikuchi 1
E-mail: juchizhang@gmail.com
Isao Takahashi 2
E-mail: takahasi@sguc.ac.jp
Saeko Tanaka 3
E-mail: saekot2356@gmail.com
Susumu Yamamoto 1
E-mail: syama@iwate-u.ac.jp
1 Department of Psychology, Iwate University, 3-18-33 Ueda, Morioka, Iwate 020-8550, Japan.
2 Department of Psychology, Sanyo Gakuen University, 1-14-1 Naka-ku Hirai, Okayama, Okayama 703-8501, Japan.
3 Department of Childhood Education, Tokushima Bunri University, Nishihama, Yamashiro, Tokushima 770-8514, Japan.

Competing Interests
There are no conflicts of interest to declare.

Citation information
Cite this article as: Is perceiving another’s error detrimental to learning from corrective feedback?, Nobuyoshi Iwaki, Mizuki Tomsawa, Reika Suzumori, Akira Kikuchi, Isao Takahashi, Saeko Tanaka & Susumu Yamamoto, Cogent Psychology (2020), 7: 1717052.

References
Barnes, J. M., & Underwood, B. J. (1959). “Fate” of first-list associations in transfer theory. Journal of Experimental Psychology, 58, 97–105. doi:10.1037/h0047507
Briggs, G. E. (1957). Retroactive inhibition as a function of the degree of original and interpolated learning. *Journal of Experimental Psychology, 53*, 60–67. doi:10.1037/h0047458

Carpenter, S. (2011). Semantic information activated during retrieval contributes to later retention: Support for the mediator effectiveness hypothesis of the testing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 1547–1552. doi:10.1037/a0024140

Delaney, P., Verkoeijen, P., & Spigel, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. *Psychology of Learning and Motivation, 53*, 63–148.

Falchikov, N. (1995). Peer feedback marking: Developing peer assessment. *Innovations in Education and Training International, 32*, 175–187. doi:10.1080/089000595012012

Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction, 20*, 304–315. doi:10.1016/j.learninstruc.2009.08.007

Grimaldi, P. J., & Karpicke, J. D. (2012). When and why do retrieval attempts enhance subsequent encoding? *Memory & Cognition, 40*, 505–513. doi:10.3758/s13421-011-0174-0

Grimaldi, P. J., & Karpicke, J. D. (2014). Guided retrieval practice of educational materials using automated scoring. *Journal of Educational Psychology, 106*, 58–68. doi:10.1037.a003208

Hovardas, T., Tsivitanidou, O. E., & Zacharia, Z. C. (2014). Peer versus expert feedback: An investigation of the quality of peer feedback among secondary school students. *Computers & Education, 71*, 133–152. doi:10.1016/j.compedu.2013.09.019

Huelsler, B. J., & Metcalfe, J. (2012). Making related errors facilitates learning, but learners do not know it. *Memory & Cognition, 40*, 514–527. doi:10.3758/s13421-011-0167-2

Izawa, C. (1970). Optimal potentiating effects and forgetting-prevention effects of tests in paired-associate learning. *Journal of Experimental Psychology, 83*, 340–344. doi:10.1037/h0028541

Jönsson, F. U., Kubik, V., Sundqvist, M. L., Todorov, I., & Izawa, C. (2014). How crucial is the response format for the testing effect? *Psychological Research, 78*, 623–633. doi:10.1007/s00426-013-0522-8

Karpicke, J. D., Lehman, M., & Aue, W. R. (2014). Retrieval-based learning: An episodic context account. *Psychology of Learning and Motivation, 61*, 237–284.

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science, 319*, 966–968. doi:10.1126/science.1152408

Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 989–998. doi:10.1037/a0015729

Kornell, N., & Vaughn, K. E. (2016). How retrieval attempts affect learning: A review and synthesis. *Psychology of Learning and Motivation, 65*, 1–33.

Lehman, M., & Karpicke, J. D. (2016). Elaborative retrieval: Do semantic mediators improve memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*, 1573–1591. doi:10.1037/xlm0000267

McDermott, K. B., Agarwal, P. K., D’Antonio, L., Roediger, H. L., III, & McDaniel, M. A. (2014). Both multiple-choice and short-answer quizzes enhance later exam performance in middle and high school classes. *Journal of Experimental Psychology: Applied, 20*, 3–21. doi:10.1037/xap0000004

Metcalfe, J., & Xu, J. (2018). Learning from one’s own errors and those of others. *Psychonomic Bulletin & Review, 25*, 402–408. doi:10.3758/s13423-017-1287-7

Postman, L., & Underwood, B. J. (1973). Critical issues in interference theory. *Memory & Cognition, 1*, 19–40. doi:10.3758/BF03198064

Putnam, A. L., & Roediger, H. L., III. (2013). Does response mode affect amount recalled or the magnitude of the testing effect? *Memory & Cognition, 41*, 36–48. doi:10.3758/s13421-012-0245-x

Pyć, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language, 60*, 437–447. doi:10.1016/j.jml.2009.01.004

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

Roediger, H. L., & Karpicke, J. D. (2006a). Test enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*, 249–255. doi:10.1111/j.1467-9280.2006.00169.x

Roediger, H. L., & Karpicke, J. D. (2006b). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1*, 181–210. doi:10.1111/j.1745-6916.2006.00012.x

Tauber, S. K., Witherby, A. E., Dunlosky, J., Rawson, K. A., Putnam, A. L., & Roediger, H. L., III. (2018). Does covert retrieval benefit learning of key-term definitions? *Journal of Applied Research in Memory and Cognition, 7*, 106–115. doi:10.1016/j.jarmc.2016.10.004

Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research, 68*, 249–276. doi:10.3102/00346543068003249

Underwood, B. J. (1957). Interference and forgetting. *Psychological Review, 64*, 49–60. doi:10.1037/h0046616

Underwood, B. J., & Freund, J. S. (1968). Effect of temporal separation of two tasks on proactive inhibition. *Journal of Experimental Psychology, 78*, 50–54. doi:10.1037/h0026157
### Appendix A. Target words and examples of false information in Experiments 1 and 2

| Target                                      | Example of false information | Exp. | Target                                      | Example of false information | Exp. |
|---------------------------------------------|------------------------------|------|---------------------------------------------|------------------------------|------|
| Brain lateralization                        | Asymmetrical ability         | E1, 2| Vicarious retribution                       | Proxy war                   | E1, 2|
| Cocktail party effect                       | Attention effect             | E1, 2| Biofeedback                                 | Bio-monitoring              | E1, 2|
| Myelin sheath                               | Synapse                      | E1, 2| Transductive reasoning                     | Transfer of reasoning       | E1, 2|
| Johari window                               | Implicit theory of personality| E1, 2| Sleeper effect                             | Time lag effect             | E1, 2|
| Zone of proximal development                | Stage of co-operation        | E1, 2| Positive illusions                         | Positive thinking           | E1, 2|
| Naive theory                                | Preconceived idea            | E1, 2| Preferential looking method                 | Eye tracking                | E1, 2|
| Halo effect                                 | Justification effect         | E1, 2| Bottom-up processing                       | Ascending processing        | E1, 2|
| Self-handicapping                           | Self-defense                 | E1, 2| Prospective memory                         | Future memory               | E1, 2|
| Boomerang effect                            | Reactance effect             | E1, 2| Spiral of silence                          | Public pressure             | E1, 2|
| Expertise                                   | Learning acceleration        | E1, 2| Prospect theory                            | Game theory                 | E1, 2|
| Door-in-the-face technique                  | Persuasive request technique | E1, 2| Face inversion effect                      | Face up-and-down effect     | E1, 2|
| Conforming behavior                         | Passive behavior             | E1, 2| Personal space                              | Private space               | E1, 2|
| Pygmalion effect                            | Adjustment effect            | E1, 2| Self-efficacy                               | Self-esteem                 | E1, 2|
| Safety base                                 | Affectionate base            | E1, 2| Cell assembly                               | Cell unit                   | E1   |
| Learning readiness                          | Maturation                   | E1, 2| Hippocampus                                 | Frontal lobe                | E1, 2|
| Critical period                             | Manifestation period         | E1, 2| Circadian rhythm                           | Solar rhythm                | E1, 2|
| Domain specificity                          | Fixability                   | E1, 2| Misattribution                              | Misreasoning                | E1, 2|
| Aptitude-treatment interaction              | Aptitude effect              | E1, 2| Mood-congruency effect                      | Mood cognition effect       | E1, 2|
| Meaningful reception learning               | Meaningful learning          | E1, 2| Self-presentation                          | Vanity gratification        | E1, 2|
| Mirror neuron                               | Mimicry cell                 | E1, 2| Develop mental acceleration                | Precocity                   | E1, 2|
| Homeostasis                                 | Physiological stabilization  | E1, 2| Placebo effect                              | Mind and body effect        | E1, 2|
| Fading                                      | Reduction                    | E1, 2| Narrative therapy                          | Reconstruction therapy      | E1   |
| Meta-reasoning                              | Recursive reasoning          | E1, 2| Masking                                    | Shielding                   | E2   |
| Counterbalancing                            | Matching                     | E1, 2| Binocular stereopsis                       | Stereo vision               | E2   |

(Continued)
| Target                | Example of false information | Exp. | Target                | Example of false information | Exp. |
|----------------------|------------------------------|------|----------------------|------------------------------|------|
| Salamanca Statement  | Inclusion Statement          | E1   | Sound image           | Stereophonic sound           | E2   |
| Motion parallax      | Binocular parallax           | E1, E2 |

Examples of statements (number of Japanese characters is in parentheses)

1. People can recognize specific contents in the noisy situation where many persons are talking. This is called the [cocktail party] effect of attention. (68)
2. There is rarely an educational method that has the same level of effect across learners. When educational treatments have different effects depending on the characteristics of learners, it is called [aptitude-treatment interaction]. (77)
3. Learners acquire knowledge and skills progressively through long-term repeated experience in a region and then show excellent performance in the region. This is called [expertise]. (70)