

Immediate Versus Delayed Feedback on Learning: Do People's Instincts Really Conflict With Reality?

Xiwen Lu
Brandeis University

Adam Sales
Worcester Polytechnic Institute

Neil T. Heffernan
Worcester Polytechnic Institute

Researchers have held differing views on the effects of feedback timing for decades. A closer reading of the timing of feedback literature that favored delayed feedback revealed that this conclusion may have been reached prematurely, because the results might have been confounded by the time interval between feedback and a posttest. This study differs from previous feedback timing studies in three distinct ways: First, this study addressed the limitations of previous studies by holding time interval between the feedback (either immediate or delayed) and the posttest constant. Second, this study included various types of knowledge and investigated the interaction between feedback timing and different knowledge types. Third, most studies that investigate the comparative effectiveness of immediate and delayed feedback on written assignments were conducted in the STEM fields, whereas few studies can be found in the second language learning field. Results revealed that the immediate feedback condition significantly outperformed the delayed feedback condition on conceptual knowledge learning, however, no difference between the two conditions was found on situational knowledge learning.

Keywords: feedback timing, immediate feedback, delayed feedback, knowledge types, second language learning

INTRODUCTION

Feedback in educational contexts has been considered crucial to learning (Narciss & Huth, 2004). Feedback effectiveness is affected by a range of factors such as the purpose, specificity, information provided, forms, as well as the timing (Shute, 2008). Among all of the above-mentioned features, the preference for immediate or delayed feedback has long been a debate. The cognitive psychology literature suggests that the timing of feedback may affect learning (Butler et al., 2007; Kulik & Kulik, 1988), and each preference has their own theories to support it (Fu & Li, 2020). Researchers have been holding opposing conclusions on the effects of feedback timing for decades (Butler et al. 2007; Clariana, 1999; Kulik & Kulik, 1988; Pound & Bailey, 1975). Most educators and students assume that immediate feedback

improves learning. There has been a significant shift toward the use of immediate feedback in online courses, and learning platforms are designed to improve learning through its use. Not everyone, however, understands the effectiveness of immediate feedback the same way. Some researchers have underscored the assumption that immediate feedback is intrinsically better as derived from the behaviorist approach to learning (Holland & Skinner, 1961). However, behaviorism is far from explaining human learning behavior based on current cognitive research results. Butler & Woodward (2018) reviewed many studies and concluded that task-level delayed feedback outperformed immediate feedback. Corral et al. (2021) conducted three experiments and argued that longer intervals in delayed feedback might enhance the processing of feedback and enhance conceptual knowledge learning. Furthermore, with the development of the cognitive revolution, many laboratory studies reported that delayed feedback can be more effective in promoting long-term retention than immediate feedback (e.g. Smith & Kimball, 2010, Mullet et al. 2014). The *ManyClasses* experiment (Fyfe et al., 2019) was a vigorous attempt in this regard and asked the question about when delayed feedback is better than immediate feedback across a variety of authentic educational contexts. They pre-registered a complete experiential analysis plan in which several instructors tested the two conditions of immediate and delayed feedback. They did not, however, find any significant difference between the two feedback conditions.

Does the desire to provide immediate feedback conflict with reality - that delayed feedback improves learning? A closer reading of the timing of feedback literature which favored delayed feedback reveals that past studies related to delayed feedback have reached premature conclusions. Many of these studies contain methodological flaws that undermine their results. For instance, the uncontrolled time interval between different kinds of feedback and posttests might have misled the results (Metcalf et al., 2009; Nakata, 2015). For example, Fyfe et al. (2019) only required a set time period between the assignment and the posttest, and designated delayed feedback provided a confounder between the assignment and the posttest, as students in the delayed feedback condition received a shorter time interval between the exposure of the question and answer to the posttest. However, the immediate feedback condition did not. Other similar experimental designs include Mullet et al. (2014). In other words, previous studies were designed inherently to support the opinion that delayed feedback is preferable for a “test with a set date”. This contrast learning in which immediate feedback is given, because the delayed feedback appears closer to the posttest, hence improving students’ test grades. The time interval confounds the results, potentially caused by the different time interval rather than feedback timing *per se*.

The types of information that students learn might also affect the feedback results, as different types of knowledge indicate various informational processing methods in the brain. Educational psychologists De Jong, T., & Ferguson-Hessler, M. (1996) summarized previous research on different types and qualities of knowledge, providing four major categories of knowledge from the practical perspective: 1. situational, 2. conceptual, 3. procedural, and 4. strategic. The first two types of knowledge, situational knowledge (such as a cultural fact) and conceptual knowledge (such as the usage of a grammar pattern), are the most commonly acknowledged in all learning fields, relatively easy to quantize, and each possess unique features. Hence, we chose the first two types of knowledge as the research subjects of the current experiment.

There have been studies that focus on the feedback timing of situational knowledge such as vocabulary learning, especially in the first language learning field. In the second language learning field, Nakata (2015) investigated the effectiveness of feedback timing on Japanese college students learning English vocabulary. They were able to hold the time interval between feedback and posttests constant to avoid confounding. However, they discovered little variance between immediate and delayed feedback on vocabulary learning. As for conceptual knowledge such as grammar patterns, it is not yet clear whether the learning of this kind of knowledge may be affected by feedback timing in the field of second language learning.

In the second language learning field, feedback for oral communication and written assignments are explored separately due to their distinct features. Most studies that focus on the corrective feedback for oral communication favored immediate feedback (Fu & Li, 2020, Li, 2020). Although there have been many studies in the STEM field on the effects of feedback timing on written assignments, few studies have been done in the second language (L2) learning field. The “feedback” mentioned in the current study referred to

written feedback, not oral corrective feedback. Nakata (2015) compared the learning effects of immediate feedback (immediately after each response) and delayed feedback (withheld until the end of the entire assignment) and found no difference between the immediate and delayed feedback. However, the delayed feedback they studied was not typical delayed feedback in a real classroom: in a traditional classroom with paper worksheets, teachers would typically not provide feedback until a few days later.

In the current study, immediate feedback refers to feedback provided immediately following each question and prior to the next one, while delayed feedback refers to feedback provided a few days after students have completed the entire assignment. This study focused on applied research in the classroom with computers rather than in laboratory research. As a result, the definition of immediate and delayed feedback underscored the difference between online assignment feedback, which can be provided right after each question, rather than the traditional method of providing assignment feedback which usually takes a few days. Feedback delay by a few seconds is considered immediate feedback, rather than delayed feedback in the laboratory (Carpenter & Vul, 2011).

This study differs from previous feedback timing studies in three keyways: First, this study addressed the limitations of previous studies by holding the time interval between the feedback (either immediate or delayed) and the posttest constant. By doing this, students in different conditions had an equal length of time to forget information, and the posttest results were not confused with a lag to posttest. Second, this study included different types of knowledge and investigated the interaction between feedback timing and different types of knowledge. Third, most studies on the comparative effectiveness of immediate and delayed feedback on written assignments were conducted in the STEM fields, and not much support can be found in the second language learning field. This study sought to fill this gap by conducting studies in a second language classroom.

This study hypothesized that immediate feedback is at least as effective as delayed feedback when the same time interval between feedback and posttest are the same. The research question is: how does feedback timing affect students' learning of situational and conceptual knowledge in a second language classroom?

METHODS

Participants

Participants were 41 undergraduate students enrolled in an Intermediate Level Chinese class in the Fall 2019 semester at a private university. The participants included 24 males and 17 females, with 3 freshmen, 20 sophomores, 13 juniors, and 5 seniors. Students participated in the course for credit.

Experimental Design

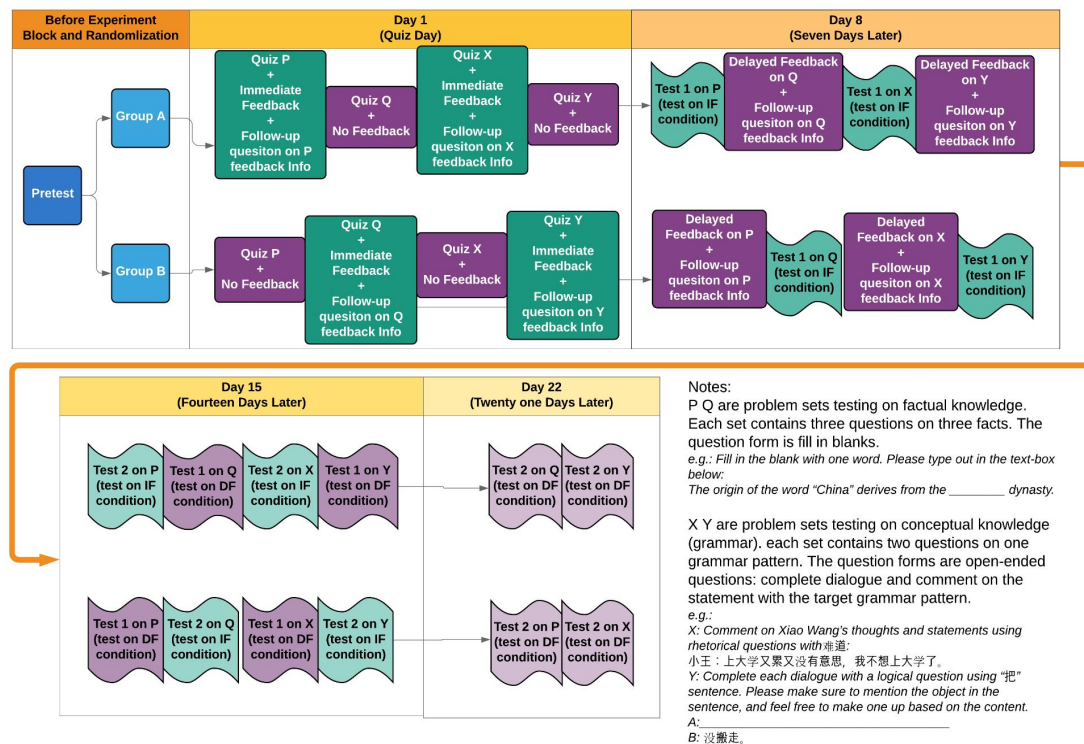
This study was a pre-post randomized experiment with a within-subject design: feedback timing (Immediate versus Delayed) was the independent variable and students' posttest results were the dependent variable. Controlled variables included test contents, pretest score, and test time.

Participants were randomly divided into two groups. To increase the chance that participants in each group had equal variance, participants were anonymized and rank-ordered based on the pretest scores. Then, the top two performing participants were paired, and each pair member was randomly assigned to a group, as shown in FIGURE 1.

Both groups received all the situational and conceptual knowledge questions in the same order of P, Q, X, Y, yet that contained different feedback conditions. Group A received question sets P and X with immediate feedback, Q and Y with delayed feedback, while Group B received question sets P and X with delayed feedback and Q and Y with immediate feedback.

All the assignments and tests took place during class time. The time interval between feedback and posttests was a period of seven days and that was determined by the course schedule.

FIGURE 1
EXPERIMENTAL DESIGN AND PROCEDURE



Materials

Assignments

Four sets of assignments were designed: P, Q, X, and Y. Quiz P and Q each contained three situational questions. Situational knowledge included a series of fun facts related to China and Chinese culture rarely known by students and challenging to recall after a single reading. Quiz X and Y each contained three conceptual questions for one grammar pattern. Two grammar patterns were implemented that students learned previously, yet had a difficult time mastering: making rhetorical questions for set X and *ba* (把) pattern for set Y. The questions for each assignment were listed in Appendix A. P and X were grouped together, Q and Y were grouped together when assigning quizzes to groups.

The Pretest and Posttests

Each test contained all the situational questions in P and Q, as well as two questions on conceptual knowledge 1 and two questions on conceptual knowledge 2. P and Q were represented situational knowledge, so the pretest and the two posttests contained the same questions on the assignments. X and Y represented conceptual knowledge so the pretest and the two posttests contained new questions about the same concepts in the assignments. The questions for all the tests can be found in Appendix B. Identical questions were used in both the pretest and posttests to better compare the learning results. Since no feedback was provided after the pretest, and the time lag between the pretest and posttest was at least 7 days, it is assumed that the pretest would not cause a significant difference in posttest learning. Target grammar patterns and cultural facts were not mentioned in the class nor in any of the assignments during the entire experimental process.

Feedback

All the situational questions could be graded automatically by the system and both types of feedback were provided: The feedback included confirmation of accuracy as well as the correct answers. For the feedback of conceptual questions, a sample answer, as well as a few sentences explaining the usage of the grammar pattern, were provided as feedback. This kind of feedback was provided because there were numerous variations for grammar pattern usage, and students' answers could not be graded automatically. All students enrolled in the course accessed and completed the assignments and received the same feedback. However, the timing of feedback varied depending on the conditions. To ensure that the feedback was read, students were required to view the feedback for each question in order to receive credit for completing the assignment. A follow-up question appeared immediately following students reading of the feedback, whether immediate or delayed, and the instructions prompted the following response: "The correct answer is ... Did you get it right?"

Procedure

This experiment was conducted during regular class hours with the use of ASSISTments embedded in the Moodle platform. ASSISTments is an online learning platform that supports student learning with hints and immediate feedback (Heffernan, N.T. & Heffernan, C.L., 2014). The assignments, immediate or delayed feedback, follow-up questions on feedback information, and the posttest were all delivered using ASSISTments. Since the course gives all participants daily in-class quizzes using student's individual laptops, all participants were comfortable taking quizzes online. Two in-class quizzes on ASSISTments were assigned prior to the experiment to help students familiarize themselves with the system.

ASSISTments anonymized the research data before providing it to researchers. All experiments were overseen by WPI's IRB team.

At the beginning of the semester, instructors announced the opportunity to participate in the study, clarifying to students that participation in the research study would not change their course experiences, but rather, stated that their course data would be provided to the researchers. Before the experiment started, students had been assigned three homework assignments on ASSISTments to familiarize themselves with the platform. A pretest was assigned during class time with no feedback given. Then, students were assigned randomly into two groups based on their pretest scores. On the first day of the experiment, the course instructor asked all students to login to the Moodle platform. Instructors provided ample class time for all students to complete their assignments. Depending on the groups, students received immediate feedback for two sets of questions and no immediate feedback for the remaining two. Immediate feedback was provided immediately following each question. If they received immediate feedback, students would be asked to answer a follow-up question to confirm that it was read.

On the eighth day, the teacher again asked all students to login to the Moodle platform and complete all the questions. Students were asked to take post-test 1 for the question sets on which they had received immediate feedback seven days ago, as well as read delayed feedback for the two sets of questions they hadn't received feedback on from seven days ago. The delayed feedback is also followed by questions to confirm the reading of the feedback. On the sixteenth day, students were required to take test 2 for the assignments with immediate feedback, and test 1 for those with delayed feedback. On the twenty-second day, students were asked to take test 2 for the assignments with delayed feedback. A procedure flow chart can be found in FIGURE 1.

Preregistration

The experimental design was pre-registered on OSF and the fully anonymized data can be found in Lu & Heffernan (2020). The sole analysis change from the plan was using R instead of SPSS. During pre-registration t-tests and ANCOVAs in SPSS were suggested, however, given the fact that we paired our students up during randomization, the stronger and more appropriate analysis accounts for that pairing. So, we began implementing R and ran the linear mixed effect model below with the pairing variable (called PairNumber) as a fixed effect.

RESULTS

One participant did not attend on the posttest days and was excluded from the study. Conceptual knowledge data and situational knowledge data were both analyzed separately.

Conceptual Knowledge X, Y

A linear mixed model fit by REML in R (RStudio version 1.3.959) is used to perform linear mixed-effects analysis of the relationship between posttest scores and conditions. T-tests used Satterthwaite's method [lmerModLmerTest]. As fixed effects, we entered the condition, test content, pretest score, and test time. As random effects, we had intercepts for subjects, and converted the variable “pair number” into a factor but preserved the variable and value label attributes. The experimental model was:

$$Score_{ij} = \beta_{0pair[i]} + \beta_1 Condition_{ij} + \beta_2 TestContent_{ij} + \beta_3 PreAveXorY_{ij} + \beta_4 TestTime_{ij} + \eta_i + \varepsilon_{ij} \quad (1)$$

where $Score_{ij}$ is the score for student i at measurement time $j=1,2,3$ (?), $\beta_{0pair[i]}$ is a fixed intercept for i 's randomization pair, and η_i is a random intercept for student i . The REML criterion at convergence was 100.1. The model's intercept remained at 0.45 (SE = 0.21, 95% CI [0.05, 0.70]). Within this model: The effect of condition was significant, $\beta_1 = 0.16$, SE = 0.04, 95% CI [0.07, 0.24], $t(110.65) = 3.71$, $p < .001^{***}$, which indicated the immediate condition performed significantly better than the delayed condition. The effect of test content was significant (beta = -0.23, SE = 0.05, 95% CI [-0.32, -0.12], $t(117.72) = -4.70$, $p < .001^{***}$). When looking into the test content, participants performed significantly better in X than in Y, which might be caused by differing question types or the varying difficulty levels of the two grammar patterns. The effect of the pretest was significant, (beta = 0.39, SE = 0.11, 95% CI [0.25, 0.68], $t(128.96) = 3.44$, $p < .001^{***}$), which indicated that higher pretest scores led to higher posttest scores. The effect of test time was insignificant, (beta = -0.001, SE = 0.04, 95% CI [-0.08, 0.09], $t(110.31) = -0.03$, $p = 0.97$). None of the effect of pair numbers was significant. Variance explained by the random effect subject was 0.03, SD = 0.19.

TABLE 1
LINEAR MIXED-EFFECTS ANALYSIS FOR CONCEPTUAL KNOWLEDGE USING POSTTEST SCORES AS THE CRITERION

Predictor	β	SE	df	t	p	95%CI
(Intercept)	0.448	0.207	39.262	2.160	0.037*	[0.053, 0.704]
Condition	0.163	0.044	110.652	3.714	0.0003 ***	[0.069, 0.241]
TestContent	-0.233	0.050	117.721	-4.704	7.01e-06 ***	[-0.315, -0.124]
Pretest Average	0.388	0.113	128.956	3.443	7.75e-04***	[0.252, 0.677]
Test Time	-0.001	0.043	110.313	-0.032	0.974	[-0.083, 0.086]

Note: $N=154$; * $p < .05$; ** $p < .001$; *** $p < .0001$.

Situational Knowledge PQ

The same model as the conceptual knowledge dataset was used for situational knowledge data. The REML criterion at convergence was -1.5. The model's intercept was at 0.28 (SE = 0.12, 95% CI [0.04, 0.41]). Within this model: The effect of condition was insignificant, beta = -0.05, SE = 0.09, 95% CI [-0.02,

0.10], $t(110.9) = -0.55$, which indicated that the condition did not explain much of the variance, and no difference was found between immediate and delayed conditions. The effect of test content was insignificant ($\beta = -0.04$, $SE = 0.03$, 95% CI [-0.03, -0.09], $t(113.4) = 1.23$, $p = 0.22$). Alternatively, the pretest effect was significant, ($\beta = 0.54$, $SE = 0.10$, 95% CI [0.38, 0.74], $t(127.7) = 5.51$, $p < .001^{***}$), which indicated that higher pretest scores led to higher posttest scores. The effect of test time was insignificant, ($\beta = 0.003$, $SE = 0.04$, 95% CI [-0.02, 0.09], $t(110.5) = 0.007$, $p = 0.99$). None of the effect of pair numbers was significant. Variance explained by the random effect “subject” was 0.02, $SD = 0.13$.

TABLE 2
LINEAR MIXED-EFFECTS ANALYSIS FOR SITUATIONAL KNOWLEDGE USING
POSTTEST SCORES AS THE CRITERION

Predictor	β	SE	df	t	p	95%CI
(Intercept)	0.232	0.123	27.629	1.888	0.070	[0.045, 0.407]
Condition	0.042	0.029	111.776	1.452	0.149	[-0.016, 0.097]
Test Content	0.038	0.031	114.430	1.233	0.220	[-0.027, 0.093]
Pretest Average	0.541	0.099	128.831	5.488	2.08e-07 ***	[0.381, 0.744]
Test Time	0.032	0.029	111.741	1.101	0.273	[-0.023, 0.090]

Note: $N=155$; $SE = \text{standard error of } \beta$; * $p < .05$; ** $p < .001$; *** $p < .0001$.

DISCUSSION

The purpose of this study is to identify the optimal feedback timing for student learning in the second language field. This study included two informational types to study the impact of feedback timing with different knowledge types. This study also kept the time interval constant, between feedback and posttests in both immediate and delayed conditions. Results of this study suggested that the immediate condition performed significantly better than the delayed condition when learning conceptual knowledge (grammar), no difference was found between the two conditions when learning situational knowledge (cultural facts).

The results of this study contradicted the findings that claim a significant delay-retention effect, and supported the effectiveness of immediate feedback when learning conceptual knowledge such as grammar. This study partially supported that immediate feedback outperformed delayed feedback by clearly defining immediate and delayed feedback in a classroom setting while holding the time interval between the feedback and posttests as constant in both conditions.

Conceptual knowledge and situational knowledge reacted differently to the feedback timing. No variation was found between the immediate and delayed feedback conditions while students were learning situational knowledge, and these results further supported Nakata (2015)’s findings. However, a significant difference was found in the two conditions in the learning of conceptual knowledge. According to Anderson (1983), conceptual knowledge is encoded declaratively first, which is identical to situational knowledge, then translated into procedures. If this is the case, then the different results on the two types of knowledge in our study can be indirect evidence that immediate feedback has a positive impact on the procedures of translating declarative information into conceptual knowledge.

In the second language learning field, most research related to immediate feedback is focused on vocabulary learning. Effects of feedback timing on the learning of grammar is not explored thoroughly. One reason might be caused by the difficulties of providing immediate feedback to open-ended questions, such as for grammar usage. For example, there are usually several correct ways to say one sentence with the same grammar pattern due to language variations. This study tries to fill this gap in the research by

providing suggested correct answers. Results indicated that the provision of correct answers improves students' online learning for open-ended questions.

It should be noted that the two different conceptual knowledge test content of X and Y made a difference in the results, and R analysis indicated that students performed significantly better in the posttest of X than in Y. In our data analysis, we discovered that X also received significantly higher pretest results compared to Y. This might be caused by varying levels of difficulty of the two grammar patterns or different question types. Test contents were selected based on the course schedule to best serve students' learning purposes. To prevent students from learning the content of the second condition while completing the first, different contents were required.

Pedagogically, the results imply that immediate feedback may be preferred when learning conceptual knowledge such as grammar in the second language field, and both delayed and immediate feedback may be used when learning situational knowledge such as cultural facts. Since most educators and students have the assumption that immediate feedback improves learning, providing immediate feedback for situational and conceptual knowledge might be more desirable. The advantage of immediate feedback also indicated that online homework assignments with feedback provided immediately after each question may be a more effective way of learning, as compared to traditional handwritten homework assignments which typically take a few days for students to receive any feedback. Although it is challenging to immediately correct open-ended questions such as forming sentences with a grammar pattern, it is nevertheless beneficial for instructors to provide immediate feedback with suggested answers.

ACKNOWLEDGEMENTS

This work was supported in part by NSF (e.g., 2118725, 1950683, 1917808, 1931523, 1940236, 1917713, 1903304, 1822830, 1759229, 1724889, 1636782, & 1535428), IES (e.g., R305D210031, R305A170137, R305A170243, R305A180401, & R305A120125), GAANN (e.g., P200A180088 & P200A150306), EIR (U411B190024), ONR (N00014-18-1-2768) and Schmidt Futures.

REFERENCES

- Anderson, J.R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 22(3), 261–295. [https://doi.org/10.1016/S0022-5371\(83\)90201-3](https://doi.org/10.1016/S0022-5371(83)90201-3)
- Butler, A., Karpicke, J., & Roediger, H. (2007). The effect of type and timing of feedback on learning from multiple-choice tests. *Journal of Experimental Psychology: Applied*, 13(4), 273–281. <https://doi.org/10.1037/1076-898X.13.4.273>
- Butler, A.C., & Woodward, N.R. (2018). Toward consilience in the use of task-level feedback to promote learning. *Psychology of Learning and Motivation*, 69. <https://doi.org/10.1016/bs.plm.2018.09.001>
- Carpenter, S.K., & Vul, E. (2011). Delaying feedback by three seconds benefits retention of face–name pairs: The role of active anticipatory processing. *Memory & Cognition* 2011, 39(7), 1211–1221. <https://doi.org/10.3758/S13421-011-0092-1>
- Clariana, R.B. (1999). *Differential Memory Effects for Immediate and Delayed Feedback: A Delta Rule Explanation of Feedback Timing Effects*. Paper presented at the Association of Educational Communications and Technology annual convention, Houston, TX.
- Corral, D., Carpenter, S.K., & Clingan-Siverly, S. (2021). The effects of immediate versus delayed feedback on complex concept learning. *Quarterly Journal of Experimental Psychology*, 74(4), 786–799. <https://doi.org/10.1177/1747021820977739>
- De Jong, T., & Ferguson-Hessler, M. (1996). Types and qualities of knowledge. *Educational Psychologist*, 31(2), 105–113.
- Fu, M., & Li, S. (2020). The effects of immediate and delayed corrective feedback on L2 development. *Studies in Second Language Acquisition*, pp. 1–33. <https://doi.org/10.1017/S0272263120000388>

- Fyfe, E., de Leeuw, J.R., Carvalho, P.F., Goldstone, R., & Motz, B. (2019, May 28). *ManyClasses 1: Assessing the generalizable effect of immediate versus delayed feedback across many college classes*. <https://doi.org/10.31234/osf.io/4mvyh>
- Heffernan, N.T., & Heffernan, C.L. (2014). The ASSISTments ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education*, 24(4), 470–497. <https://doi.org/10.1007/s40593-014-0024-x>
- Holland, J.G., & Skinner, B.F. (1961). *The analysis of behavior: A program for self-instruction*. New York, NY, US: McGraw-Hill.
- Kulik, J.A., & Kulik, C-L.C. (1988). Timing of Feedback and Verbal Learning. *Review of Educational Research*, 58(1), 79–97. <https://doi.org/10.3102/Timing of Feedback and Verbal Learning - James A. Kulik, Chen-Lin C. Kulik, 1988/00346543058001079>
- Li, S. (2020). What is the ideal time to provide corrective feedback? Replication of Li, Zhu & Ellis (2016) and Arroyo & Yilmaz (2018). *Language Teaching*, 53(1), 96–108. <https://doi.org/10.1017/S026144481800040X>
- Lu, X., & Heffernan, N.T. (2020, July 14). *Data*. Retrieved from osf.io/wvfxk
- Metcalf, J., Kornell, N., & Finn, B. (2009). Delayed versus immediate feedback in children’s and adults’ vocabulary learning. *Memory & Cognition*, 37, 1077–1087. <https://doi.org/10.3758/MC.37.8.1077>
- Mullet, H.G., Butler, A.C., Verdin, B., von Borries, R., & Marsh, E.J. (2014). Delaying feedback promotes transfer of knowledge despite student preferences to receive feedback immediately. *Journal of Applied Research in Memory and Cognition*, 3, 222–229. <https://doi.org/10.1016/j.jarmac.2014.05.001>
- Narciss, S., & Huth, K. (2004). How to design informative tutoring feedback for multi-design for multimedia learning. In H. M. Niegemann, D. Leutner, & R. Brunken (Ed.), *Instructional design for multimedia learning* (pp. 181–195). Munster, NY: Waxmann.
- Nakata, T. (2015). Effects of feedback timing on second language vocabulary learning: Does delaying feedback increase learning? *Language Teaching Research*, 19(4), 416–434. <https://doi.org/10.1177/1362168814541721>
- Pound, L.D., & Bailey, G.D. (1975). Immediate Feedback Less Effective than Delayed Feedback for Contextual Learning. *Reading Improvement*, 12(4), 222–224.
- Shute, V.J. (2008). Focus on Formative Feedback. *Review of Educational Research*, 78(1), 153–189. <https://doi.org/10.3102/0034654307313795>
- Smith, T.A., & Kimball, D.R. (2010). Learning from feedback: Spacing and the delay–retention effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(1), 80–95. <https://doi.org/10.1037/a0017407>

APPENDIX A: ASSIGNMENT QUESTIONS FOR P, Q, X, Y

Assignment P:

1. The origin of the word “China” derives from the Qin dynasty.
2. China is considered to be the oldest civilization with some historians marking 6000 BC as the beginning of the Chinese civilization. Also, it has the world’s longest-used written language.
3. The full name of the current leader/Chairman of China is Xi Jinping.

Assignment Q:

1. The population of China now is around 1.4 billion.
2. China has been the source of many innovations, scientific discoveries and inventions. This includes the Four Great Inventions: papermaking, the compass, gunpowder, and printing.
3. The people's republic of china was founded in 1949.

Assignment X:

Comment on Xiao Wang’s thoughts and statements using rhetorical questions with 难道:

1. 小王: 我知道明天有考试, 可是今晚我还是要去看电影。

Your comment:

A sample answer: 难道你复习好了吗?

The subject “你” can be put in front of “难道” or right after it.

Did you get it right?

2. 小王: 我三天没有睡觉了。

Your comment:

A sample answer: 难道你不累吗?

Generally speaking, rhetorical questions that are affirmative in form carry an emphatically negative meaning; rhetorical questions that are negative in form carry an emphatically positive meaning.

The meaning of the sample answer is: “Don’t you feel tired?”

Did you get it right?

3. You: 你的钱包丢了, 这很糟糕。但你的钱包不一定是小张偷的。

小王: 肯定是小张偷走的!

You:

A sample answer: 你怎么知道? 难道你看见了吗?

The underline meaning of the sample answer is “Since you didn’t see it by yourself, you cannot say that Xiao Zhang stole it.”

Did you get it right?

Assignment Y:

Complete each dialogue with a logical question using “把” sentence. Please make sure to mention the object in the sentence, and feel free to make one up based on the content.

1. A: _____
B: 写完了。

Sample answer: 你把作业写完了吗?

The object should be known. You have to make one up which goes with the verb.

把字句 are most often used to describe what happened to the object in some detail.

Remember to use “了” as in this dialogue the action has taken place.

The verb is not just “bare”; there’s “more stuff” after it. In this case, a resultative complement and “了” is needed.

Did you get it right?

2. A: _____
 B: 冰箱里。
 Sample answer: 你把蛋糕放在哪儿了?
 把 sentences are most often used to describe location change of the object. When talking about location, you should think about 在 in front of the question word. And if it has happened, “了” is also needed.
 Did you get it right?
3. A: _____
 B: 送给妈妈。
 Sample answer: 你想把这个礼物送给谁?
 When there are two objects in one sentence, 把 sentence is often used.
 把字句 are not tied to any particular time. When talking about plans, you do not need “了” in the question.
 Did you get it right?

APPENDIX B: PRETEST AND POSTTEST QUESTIONS FOR P, Q, X, Y

Tests for P (same to the assignment questions)

4. The origin of the word “China” derives from the Qin dynasty.
5. China is considered to be the oldest civilization with some historians marking 6000 BC as the beginning of the Chinese civilization. Also, it has the world’s longest-used written language.
6. The full name of the current leader/Chairman of China is Xi Jinping.

Tests for Q (same to the assignment questions)

4. The population of China now is around 1.4 billion.
5. China has been the source of many innovations, scientific discoveries and inventions. This includes the Four Great Inventions: papermaking, the compass, gunpowder, and printing.
6. The people's republic of china was founded in 1949.

Tests for X (different from the assignment questions)

Comment on Xiao Wang’s thoughts and statements using rhetorical questions with 难道:

1. 小王：上大学又累又没有意思，我不想上大学了。

Your comment:

2. 小王：我的老板很糟糕，我每天工作时心情都不好。

Your comment:

Tests for Y (different from the assignment questions)

Complete each dialogue with a logical question using “把” sentence. Please make sure to mention the object in the sentence, and feel free to make one up based on the content.

1.

A: _____

B: 没搬走。

2.

A: _____

B: 卖给了一个学生。