IESD Research Paper

What Research Says About Promoting Learning Through Retrieval Practice with Formative Feedback

August 2020

Prepared for UWorld, LLC
By Interactive Educational Systems Design, Inc.
New York, NY
Introduction

In a typical year, standardized college readiness exams play an important role in college and university admissions. More than 2.2 million students in the 2019 graduating class took the SAT® exam, and almost 1.8 million took the ACT® exam (College Board, 2019; National ACT, 2019). In addition, more than 1.2 million high school students took an Advanced Placement® (AP®) exam in their attempt to obtain college credit upon admission (Anderson, 2020).

Besides high-quality classroom instruction and core curriculum materials, high school students need effective supplemental college preparation resources that give them the greatest opportunity to succeed on these high-stakes exams.

UWorld provides research-based, college readiness digital tools for the PSAT/NMSQT®, SAT, ACT, and a variety of AP exams. UWorld digital tools feature questions and problems that mirror those on the actual exams; helpful, explanatory feedback to reinforce what students have already learned and address misconceptions and gaps in learning; and a reporting system that identifies where students should focus their attention in subsequent study and practice, tailored to each student’s individual needs.

This paper presents research and expert opinion on providing practice with elaborated, explanatory feedback to promote recall and enhance learning. The paper goes on to explain how UWorld college readiness tools align with the research literature.

---

1 It is acknowledged that 2020 is an atypical year due to the coronavirus pandemic and temporary cancellations of college readiness exams.
Executive Summary of Research on Retrieval Practice Combined with Formative Feedback

Providing formative feedback is essential. Decades of research provide evidence that high quality formative feedback improves learning and skill development, as well as student motivation to learn (Shute, 2008, citing multiple sources). Feedback has been found to be effective across a range of subject areas (Swart, Nielen, & Sikkema-de Jong, 2019; Van der Kleij, Feskens, & Eggen, 2015).

Retrieval practice aids recall. Empirical research in cognitive science supports teaching strategies focused on “retrieval practice,” that is, student attempts to recall and apply previously learned knowledge or skills after a period of forgetting. In the context of retrieval practice, feedback improves the learners’ metacognition—their understanding of what they know and do not know (Agarwal & Bain, 2019).

Elaborated feedback is most effective. Research has found that the most effective type of feedback is “elaborated” feedback that provides students with an explanation for a correct or incorrect answer. This has been shown to be more effective than simply providing an indication of whether the student’s response was correct or identifying the correct answer (Shute, 2008; Shute & Rahimi, 2017; Van der Kleij et al., 2015). However, providing the correct answer is also considered to be beneficial to learning (Agarwal & Bain, 2019; Swart et al., 2019).

Research supports providing both immediate and delayed feedback. Immediate feedback has been shown to improve retention, while delayed feedback can support transfer of learning to new situations (Agarwal & Bain, 2019; Dihoff, Brosvic, Epstein, & Cook, 2004; Shute, 2008; Swart et al., 2019; Van der Kleij et al., 2015).

Computer-based learning systems that provide practice with elaborated feedback are effective alternatives to human tutoring (Swart, et al., 2019; Van der Kleij et al., 2015).

UWorld computer-based test readiness tools provide practice questions similar to ones appearing on the actual college readiness exams; timely elaborated, explanatory feedback; flexibility in providing immediate or delayed feedback; and reporting features that help students and teachers customize subsequent practice assignments to meet the individual needs of each student.
Research on Retrieval Practice Combined with Formative Feedback

There is consensus in the education research community that feedback is an essential component of the learning process. In her seminal review of research on feedback in education, Shute (2008) summarized decades of evidence that well-designed formative feedback—that is, “information communicated to the learner that is intended to modify the learner's thinking or behavior”—improves learning and skill development (Shute, 2008, p. 153, citing multiple sources). Such feedback can also improve student motivation to learn (pp. 153, 157, 161, 163, 166, 172, citing multiple sources).

Positive Effects Across Subject Areas

Research on feedback has found positive effects across a wide range of subject areas.

A 2019 meta-analysis of 60 studies involving over 6,000 students (kindergarten through college age) working on improving reading comprehension skills compared the learning impacts of reading with and without feedback. The research team found a “robust” effect of feedback on learning (Swart, Nielen, & Sikkema-de Jong, 2019, section 3.2).

Other researchers studying feedback have found positive impacts on learning in the sciences, social sciences, math, and history (Van der Kliej, Feskens, & Eggen, 2015, p. 502; Shute & Rahimi, 2017, p. 9).

Retrieval Practice Combined with Feedback

In their book Powerful Teaching, Agarwal and Bain (2019) draw on empirical research in cognitive science to develop recommendations on effective teaching strategies focused on “retrieval practice.” Retrieval practice occurs in activities where “learners recall and apply multiple examples of previously learned knowledge or skills after a period of forgetting” (Agarwal & Bain, 2019, p. 37, citing Lemov, 2017).

Agarwal and Bain conclude that giving students opportunities for retrieval practice combined with feedback is effective in promoting learning and retention. The retrieval practice “boosts learning by pulling information out of students’ heads… rather than cramming information into students heads (e.g. lectures)” (p. 4). Feedback incorporated into retrieval practice enhances learning by providing students with “the opportunity to know what they know, and know what they don’t know. This increases students’ metacognition or understanding about their own learning progress” (p. 5).

Agarwal and Bain explain that when retrieval practice is combined with feedback, the feedback counterbalances the learner’s tendency toward overconfidence—thinking they remember or understand when, in actuality, they do not (p. 126). They cite their own classroom research and other researchers confirming that this kind of retrieval practice—giving students the opportunity

---

2 Details on guidance for high quality formative feedback are presented later in this paper.
3 Average effect was $g_{+} = 0.35 (k = 104, SE = 0.05, 95\% CI = [0.25, 0.46], p < .001)$ (Swart et al, 2019, section 3.2).
4 Swart et al. (2019) is an online-only source and does not provide page numbers. Instead, in-text citations of Swart et al. throughout this report refer to section numbers or, in some case, the abstract.
to make mistakes and then correct them—improves learning (Agarwal & Bain, 2019, pp. 136-137, citing Metcalfe & Finn, 2012; Nietfeld, Cao, & Osborne, 2006)

**Elaborated Feedback and Correct Response Feedback**

Education research strongly supports the efficacy of incorporating *elaborated* feedback—providing students with an explanation for a correct or incorrect answer—as part of practice.

For example, Van der Kliej et al. (2015) conducted a meta-analysis that compared student learning outcomes from different methods of feedback, in a computer-based learning environment (p. 499). These researchers found that “EF [Elaborated Feedback] was more effective than KR [Knowledge of Results] and KCR [Knowledge of Correct Response]” (p. 501). Their results also showed that elaborated feedback had a greater impact on higher order learning than lower order learning (p. 501).

Shute’s (2008) review of the research literature on feedback identified multiple studies where elaborated feedback was found to be “more effective than verification of results” (p. 177, citing Bangert-Drowsn et al., 1991; Gilman, 1969; Mason & Bruning, 2001; Narciss & Huth, 2004). This advantage for elaborated feedback was confirmed in a more recent meta-analysis by Shute & Rahimi (2017, p. 15).

Other researchers have found that students benefit from both elaborated and correct answer feedback (Agarwal & Bain, 2019, p. 134; Swart et al., 2019, section 5). Swart et al. found evidence of this in their meta-analysis specifically focused on reading comprehension skills (2019, section 5).

**Guidance on High Quality Formative Feedback**

Based on her review of the research literature, Shute (2008) offered guidance on effective feedback, including the following:

- Elaborated feedback can address the topic, the response, or the particular error(s). It should “describe the what, how, and why of a given problem” (pp., 158, 177).
- Provide feedback in “manageable units” that are “specific and clear” (p. 177).
- Clarify learning goal(s), “how well learners are performing on a task,” and “what needs to be accomplished” to reach the goal(s) (p. 177).
- Give unbiased, objective feedback. Computer-based feedback is perceived as unbiased (p. 177).
- Provide feedback after learners have attempted a solution (p. 178).

---

5 “The mean weighted overall effect size for EF was 0.49, which can be considered a moderately large effect. …The weighted overall mean effect size of EF contrasted with no feedback was… 0.61. The mean weighted overall effect size for KCR (k = 9) was 0.32, which is considered small to moderate. The mean weighted effect size for KR (k = 8) was extremely small, at 0.05” (Van der Kleij et al., 2015, p. 501).

6 For higher order learning outcomes, ES’ = 0.67; for lower order learning outcomes, ES’ = 0.37 (Van der Kleij et al., 2015, p. 501).
**Timing of Feedback**

There is a robust body of research indicating that immediate feedback results in improved retention and academic achievement. In their comprehensive review of the research on immediate feedback, Dihoff, Brosvic, Epstein, and Cook (2004) found that immediate feedback provided a learning advantage for students at the elementary, middle school, high school, and college levels (p. 230).

In their own research, Dihoff et al. conducted two studies in which the availability of immediate feedback and the timing of the feedback were varied during practice tests in preparation for a multiple-choice final exam. The results demonstrated that “immediate feedback increases retention, accuracy at identifying initial responses, and confidence in response identifications while simultaneously reducing [repeated] incorrect responding” (Dihoff et al., 2004, p. 230).

In Swart et al.’s meta-analysis of effects of feedback on improving reading comprehension skills, these researchers found that “feedback is particularly effective if provided directly after reading, but less so when provided during reading” (Swart, et al., 2019, abstract).

Van der Kliej et al. (2015) cited research indicating that “students prefer immediate to delayed feedback” and that they spend “significantly more time reading immediate feedback than delayed feedback” (pp. 478-479, citing Miller, 2009; Van der Kleij, Eggen, Timmers, & Veldkamp, 2012).

**Benefits of Both Immediate and Delayed Feedback**

Other research suggests that both immediate and delayed feedback are effective.

Cognitive research summarized by Agarwal and Bain indicates that both types of feedback deliver positive results (2019, p. 135).

Shute, in her review of research on formative feedback, found that immediate feedback is appropriate for verbal and procedural skills, and that it “can help fix errors in real time” and “provides a helpful safety net” when “learning a difficult new task” (2008, p. 179, citing multiple sources). She found evidence that delayed feedback is appropriate when learning a simple task and may be more effective for transfer of learning to new situations (p. 179, citing multiple sources).

**Computer-Based Learning Systems that Provide Practice with Elaborated Feedback**

Over the past few decades, researchers have focused attention on computer-based environments for providing feedback to increase learning, as an alternative to one-to-one human tutoring. There is a strong body of research that supports the effectiveness of human tutoring because the tutor can adapt instruction “continuously to the needs of the learner” (Van der Kleij et al., 2015, p. 479, citing Bloom, 1984). However, human tutoring is not feasible as a key instructional method in K-12 education.
In their meta-analysis, Van der Kleij et al. concluded that well-designed computer-based systems can help meet the need for individualized learning. These researchers referred to systems that include practice questions, “timely feedback,” and “automated scoring” (2015, p. 479).

The meta-analysis conducted by Swart et al. found that computer-delivered feedback was more beneficial for learning reading comprehension skills, compared to non-computer-delivered feedback. They advised that when considering computer-based learning systems for comprehension skill development, it is recommended to choose a system with questions and elaborated feedback directly after reading, “to help the reader evaluate and, if necessary, revise the mental model of a text” (2019, section 5).

As presented above, Shute (2008) notes that an advantage of computer-based systems is that they deliver feedback that is perceived by students as unbiased and objective (p. 177), so they are likely to engage with the feedback.
How UWorld College Readiness Tools Align with the Research

UWorld technology-enhanced college readiness tools for the PSAT, SAT, ACT, and AP course exams are designed to enhance retrieval practice with a variety of elaborated, explanatory feedback; automated scoring; and detailed reporting to enable students and teachers to determine next steps in continuing their preparation for these important exams.

**UWorld college readiness tools are designed for effective retrieval practice.**
UWorld tools are designed as practice test supplements to classroom instruction, to help students retrieve facts, concepts, and skills previously taught, and help them apply this learning to new contexts. Users select a subject or specific topic they want to “study” and then work through a set of questions on that subject or topic. Informative feedback helps reinforce previously learned content and enables students to address misunderstandings and avoid mistakes in the future.

**UWorld college readiness tools provide students with practice in the various subject areas cited in the research literature.**
For example, the UWorld SAT tool covers all SAT topics in Reading, Grammar, and Math. The ACT tool addresses all ACT topics in Reading, Grammar, Math, and Science.

There are UWorld AP tools for a variety of courses including:
- Language and Composition
- Literature
- US History
- World History
- Calculus AB
- Calculus BC
- Statistics
- Chemistry
- Physics
- Psychology
- Biology
UWorld tools to support additional AP courses are in development.

**UWorld college readiness tools feature a variety of helpful, elaborated feedback.**
Where appropriate, the feedback addresses the topic, reproducing and explaining the language of the question to help students identify the *what* of the question.

UWorld feedback is specific to each response choice. Explanations provide a concise yet detailed explanation of the correct answer that provides background on the question content, connections to the prompt or question text, and additional context where needed. This type of feedback enables students to quickly identify the *how* and the *why* of the correct answer choice.

Sometimes the feedback addresses the specific error made by the student. This type of feedback helps students fill in gaps in their knowledge or correct misunderstandings.

The *Things to remember* feature provides students with a final note about the most important information to remember about the topic.
**UWorld follows research-based guidance on high-quality informative feedback.**

UWorld always engages students in an **interactive process**, in which they first attempt to answer a question (or several questions), and then receive helpful, informative feedback.

As described above, UWorld feedback addresses the topic of the question, the correct response, and/or specific errors. It explains the **what**, **how**, and **why** of the problem.

Explanations in UWorld feedback always come in **manageable units**. The explanation of the correct response is organized into a few short paragraphs. Feedback for an incorrect response is usually a single sentence. The **Things to remember** feature restates the most important takeaways from the explanation in a sentence or two. The attention to limiting word counts, the use of **clear, concise language specific to the topic of the question**, and the use of boldface or underlining to clearly identify the correct answer all help ensure that feedback messages are easy to digest and never overwhelming.

Feedback is always presented in an **objective** manner, and students regard UWorld computer-based learning tools as providing **unbiased, trustworthy** information. Consequently, when informative feedback is offered to students through UWorld tools, they are likely to consider it and learn from it.

UWorld performance reports provide students with the data necessary to **determine how well they are doing** overall and on specific curriculum units.

The next several pages present examples from UWorld college readiness tools, with research-based feedback features explained.
Example 1. AP English Literature

The Question
The AP English Literature exam requires students to read a passage excerpt. This question references *The Blithedale Romance* by Nathaniel Hawthorne.

Which of the following most aptly describes Priscilla’s interactions with the narrator?

- **A** She cheerily dismisses both his pessimism and rational philosophy.
- **B** She inspires him by her display of loyalty to her current companions.
- **C** She surprises him with her refusal to think seriously about her behavior.
- **D** She chides him until he finally accepts her view of the world.
- **E** She stuns him with her forceful words and tone.
The Informative Feedback

To describe Priscilla's interactions with the narrator, consider their conversation—what each says and how each says it. The correct answer has the most supporting evidence.

<table>
<thead>
<tr>
<th>What the narrator says</th>
<th>What Priscilla says</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;I love to see sufficient cause for everything...&quot;</td>
<td>&quot;I never think about it at all...&quot;</td>
</tr>
<tr>
<td>He cannot understand what makes Priscilla so</td>
<td>Priscilla rejects his intellectual analysis of</td>
</tr>
<tr>
<td>happy and questions her reason.</td>
<td>happiness; &quot;laughing.&quot;</td>
</tr>
<tr>
<td>&quot;...our friends of today will not be our friends of a</td>
<td>&quot;Not ill...I will live and die with these!&quot;</td>
</tr>
<tr>
<td>few years hence....&quot;</td>
<td>Priscilla rejects his bleak forecast by quickly</td>
</tr>
<tr>
<td>He makes a gloomy prediction of future friendships.</td>
<td>affirming her devotion to her friends.</td>
</tr>
</tbody>
</table>

By questioning Priscilla and predicting gloom, the narrator demonstrates his pessimism (negativity) and rational philosophy (perspective based on logic). Preferring a positive, emotional approach to life, Priscilla lightly dismisses (rejects) his point of view, therefore, she cheerily dismisses both his pessimism and rational philosophy.

(Choice B) Although Priscilla displays loyalty to her current companions, her interaction doesn't inspire the narrator. He maintains his pessimism about the fickleness (changeability) of friendship.

(Choice C) Even though Priscilla doesn't seem to think seriously about her behavior, the narrator gives no evidence that he is surprised by this response.

(Choice D) Priscilla does not chide (scold) the narrator for his ideas, nor does he accept her view of the world. At the conclusion of their interactions, the narrator continues to feel gloomy about people and life.

(Choice E) There is no evidence that Priscilla's responses to the narrator are forceful (powerful). The narrator's view of the world remains unchanged, and he is not stunned by Priscilla's words or tone.

Things to remember:
Look at what a character says and does to find supporting evidence for a description of the character's interactions.

Features of the Feedback

- The feedback begins by reproducing key language from the question, “Priscilla's interactions with the narrator,” and explains the what of the question—that it is essentially asking which answer choice best summarizes the details of a conversation between two characters.
- The explanation of the correct choice uses a diagram to help break up the feedback into manageable chunks and makes it easy to focus on the most relevant information (textual evidence).
- The feedback includes concise explanations of each wrong answer that reinforce that the correct answer choice should be based on evidence from the text.
- The Things to remember at the end of the feedback combines the critical elements of the comprehension strategy in a clear, concise summary sentence.
Example 2. AP US History

The Question

The cartoon was most likely intended to do which of the following?

A. Address treaties that gave European nations exclusive rights to import goods from China
B. Dispel concerns about U.S. tariffs on Chinese imports
C. Raise questions about the role of the United States as global mediators
D. Assert the United States’ right to trade with China

Answer: D
The Informative Feedback

The emergence of the United States as an industrial powerhouse in the late 19th century increased desire to export American goods to overseas markets, particularly Asia. The problem however, was that the United States was relatively late in expanding export markets in the East. By 1899 Italy, England, France, Germany, Russia, and Japan had each established treaties that created spheres of influence, coastal regions where each country had exclusive trading rights.

The intent of the cartoon is to assert the United States' right to trade with China, despite previous European trade treaties. In the era of American imperialism, securing overseas markets for American exports dictated U.S. foreign policy. Of particular importance was breaking the grip that European powers held on Asian markets. In order to not be 'cut out' from Chinese markets, US Secretary of State John Hay proposed the Open Door Policy (1899). The policy asserted that nations should be allowed equal access to all markets in China.

(Choice A) Treaties that gave European nations exclusive rights to export to China—rather than import from China—are the subject of cartoon.
(Choice B) Exports to China, rather than tariffs on US imports from China, is the subject of the cartoon.
(Choice C) Rather than mediating conflict (act as a peace maker), the cartoon portrays America's determination to move into Chinese markets. By the turn of the 20th century, the US hadn't yet assumed the role of international mediator.

Things to remember:
By the turn of the century, America's industrial age resulted in innovations in production and a booming economy. US foreign policy, particularly the Open Door Policy (1899), aimed to expand the US export market into China.

Features of the Feedback

- Annotations included with the images are used to provide context while keeping the explanation concise. For example, here the annotations identify historical figures that might be unrecognizable to a modern audience. They also help explain how the image should be interpreted.

- The remainder of the correct answer explanation presents historical background context to clarify how and why the image relates to the right answer. The sentence containing the phrase in bold type connects the explanation to the language of the question (“the intent of the cartoon...”) and to the correct answer choice.

- Blue text in the explanation of the correct choice signals hyperlinks to images providing additional context for students unfamiliar with the terms—without lengthening the top-level explanation or interrupting the flow for students who already understand the terms.
Wrong answer choices address potential mistakes: misreading of “import” as “export” or misremembering the direction of trade related to the treaties and tariffs.

**Example 3. SAT Math**

**The Question**

An office supply store prints documents that have a front cover and a back cover. The store charges $0.20 for each page and $1.00 each for the front cover and the back cover when printing these documents. Which of the following gives the total cost \( f(p) \), in dollars, that the store charges to print a document that has \( p \) pages?

- (A) \( f(p) = 1.00 + 0.20p \)
- (B) \( f(p) = 2.00 + 0.20p \)
- (C) \( f(p) = 0.20 + 0.20p \)
- (D) \( f(p) = 0.40 + 0.20p \)

**The Informative Feedback**

The total cost to print a document is the sum of the fixed cost of the front and back covers and the variable cost of \( p \) pages. The variable cost is the cost per page multiplied by the number of pages.

Write an equation for the total cost \( f(p) \), then plug in the given values for the fixed cost ($1.00 each for the front cover and back cover), the cost per page ($0.20), and the number of pages \( p \).

\[
\begin{align*}
\text{total cost} &= \text{fixed cost} + \text{variable cost} \\
&= (\text{front cover} + \text{back cover}) + (\text{cost per page} \times \text{number of pages}) \\
&= (1 + 1) + (0.20)(p) \\
f(p) &= 2 + 0.20p
\end{align*}
\]

Simplify the equation to see that the total cost, in dollars, that the store charges to print a document with \( p \) pages is \( f(p) = 2.00 + 0.20p \).

- (Choice A) \( f(p) = 1.00 + 0.20p \) may result from the misconception that the front and back covers have a total cost of $1.00, but it is given that each cover costs $1.00 for a total cost of $2.00.

- (Choices C and D) \( f(p) = 0.20 + 0.20p \) and \( f(p) = 0.40 + 0.20p \) may result from incorrectly calculating the fixed cost by multiplying the cost of the front or back cover by the cost of each page ($0.20). However, the fixed cost is the sum of the cost of the front cover and the cost of the back cover ($1.00 + $1.00 = $2.00).

**Things to remember:**

- A total amount (ex. cost, time) is generally the sum of a fixed amount (ex. initial amount) plus any variable amount (ex. number of items multiplied by the rate per item).

**Features of the Feedback:**

- The color-coded visual representation of mathematical problem-solving helps keep the explanation clear and concise.
- Explanations of wrong answer choices address common potential mistakes, clarifying at what point in the solution process a mistake might have yielded each wrong answer choice.
The Things to remember section summarizes the what and how of the question. This section also serves as a generalizable approach to a type of algebraic problem.

Example 4. ACT Science

The Question

The number of body bends was counted during the 10 minute period for all groups. The resulting data are shown in Figure 1.

Consider the total number of body bends for the TG nematodes with exercise group on Day 1. Based on that result, over the 10 minutes the average body bends per minute would have been closest to which of the following?

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>5 body bends/minute</td>
</tr>
<tr>
<td>✔ B</td>
<td>50 body bends/minute</td>
</tr>
<tr>
<td>C</td>
<td>100 body bends/minute</td>
</tr>
<tr>
<td>D</td>
<td>500 body bends/minute</td>
</tr>
</tbody>
</table>
The Informative Feedback

Features of the Feedback

- The annotated graph contributes to the clear, concise explanation. The key steps of the solution process are summarized in just a few words.
- The text explanation breaks down the how of the solution process into manageable steps.
- Language throughout the explanation is both authoritative and objective, providing the steps necessary to answer the question—without any hint of negativity toward a student who might have selected a wrong answer.
**Example 5. SAT Writing Score Report**

**Features of the Feedback**

- Overall score indicates that the student’s performance on Agreement is acceptable, but individual sub-section scores reveal areas of weakness: Comparisons and Modifiers. This information can help the teacher and student target areas that require reteaching, remediation, and additional practice.

- The analysis also reveals several areas of strength where performance was less than perfect. These might be targets of further review.
**UWorld college readiness tools offer a choice between immediate and delayed feedback.**

When instructors or students are setting up a practice test, they are prompted to select whether explanatory feedback will be provided immediately after answering each question or delayed until after all questions on the test are answered. This decision to provide immediate or delayed feedback can be made on a case-to-case basis, consistent with recommendations by Shute (2008).

- After a teacher has just introduced a new content topic or skill, he or she can choose to provide immediate feedback after each question.
- Alternatively, if a class has covered several units or topics extensively but has had limited practice in formal testing situations, the teacher can choose to provide delayed feedback, after the completed practice exam has been submitted—to stimulate transfer of learning.

**UWorld college readiness tools are well-designed computer-based systems.**

UWorld computer-based test readiness systems provide practice questions similar to ones appearing on the actual college readiness exams; timely elaborated, explanatory feedback; flexibility in providing immediate or delayed feedback; and reporting features that help students and teachers customize subsequent practice assignments to meet the individual needs of each student.
References


