### Table 1. Institute of Education Sciences Levels of Evidence

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strong</strong></td>
<td>In general, characterization of the evidence for a recommendation as strong requires both studies with high internal validity (i.e., studies whose designs can support causal conclusions), as well as studies with high external validity (i.e., studies that in total include enough of the range of participants and settings on which the recommendation is focused to support the conclusion that the results can be generalized to those participants and settings). Strong evidence for this practice guide is operationalized as:</td>
</tr>
<tr>
<td></td>
<td>• A systematic review of research that generally meets the standards of the What Works Clearinghouse (see <a href="http://ies.ed.gov/ncee/wwc/">http://ies.ed.gov/ncee/wwc/</a>) and supports the effectiveness of a program, practice, or approach, with no contradictory evidence of similar quality; OR</td>
</tr>
<tr>
<td></td>
<td>• Several well-designed, randomized, controlled trials or well-designed quasi-experiments that generally meet the standards of the What Works Clearinghouse and support the effectiveness of a program, practice, or approach, with no contradictory evidence of similar quality; OR</td>
</tr>
<tr>
<td></td>
<td>• One large, well-designed, randomized, controlled, multisite trial that meets the standards of the What Works Clearinghouse and supports the effectiveness of a program, practice, or approach, with no contradictory evidence of similar quality; OR</td>
</tr>
<tr>
<td></td>
<td>• For assessments, evidence of reliability and validity that meets The Standards for Educational and Psychological Testing.</td>
</tr>
<tr>
<td><strong>Moderate</strong></td>
<td>In general, characterization of the evidence for a recommendation as moderate requires studies with high internal validity but moderate external validity, or studies with high external validity but moderate internal validity. In other words, moderate evidence is derived from studies that support strong causal conclusions but where generalization is uncertain, or studies that support the generality of a relationship but where the causality is uncertain. Moderate evidence for this practice guide is operationalized as:</td>
</tr>
<tr>
<td></td>
<td>• Experiments or quasi-experiments generally meeting the standards of the What Works Clearinghouse and supporting the effectiveness of a program, practice, or approach with small sample sizes and/or other conditions of implementation or analysis that limit generalizability, and no contrary evidence; OR</td>
</tr>
<tr>
<td></td>
<td>• Comparison group studies that do not demonstrate equivalence of groups at pretest and therefore do not meet the standards of the What Works Clearinghouse but that (a) consistently show enhanced outcomes for participants experiencing a particular program, practice, or approach and (b) have no major flaws related to internal validity other than lack of demonstrated equivalence at pretest (e.g., only one teacher or one class per condition, unequal amounts of instructional time, highly biased outcome measures); OR</td>
</tr>
<tr>
<td></td>
<td>• Correlational research with strong statistical controls for selection bias and for discerning influence of endogenous factors and no contrary evidence; OR</td>
</tr>
<tr>
<td></td>
<td>• For assessments, evidence of reliability that meets The Standards for Educational and Psychological Testing but with evidence of validity from samples not adequately representative of the population on which the recommendation is focused.</td>
</tr>
<tr>
<td><strong>Low</strong></td>
<td>In general, characterization of the evidence for a recommendation as low means that the recommendation is based on expert opinion derived from strong findings or theories in related areas and/or expert opinion buttressed by direct evidence that does not rise to the moderate or strong levels. Low evidence is operationalized as evidence not meeting the standards for the moderate or high levels.</td>
</tr>
</tbody>
</table>

---

3 Ibid.
<table>
<thead>
<tr>
<th>Recommendation</th>
<th>Level of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Space learning over time. Arrange to review key elements of course content after a delay of several weeks to several months after initial presentation.</td>
<td>Moderate</td>
</tr>
<tr>
<td>2. Interleave worked example solutions with problem-solving exercises. Have students alternate between reading already worked solutions and trying to solve problems on their own.</td>
<td>Moderate</td>
</tr>
<tr>
<td>3. Combine graphics with verbal descriptions. Combine graphical presentations (e.g., graphs, figures) that illustrate key processes and procedures with verbal descriptions.</td>
<td>Moderate</td>
</tr>
<tr>
<td>4. Connect and integrate abstract and concrete representations of concepts. Connect and integrate abstract representations of a concept with concrete representations of the same concept.</td>
<td>Moderate</td>
</tr>
<tr>
<td>5. Use quizzing to promote learning. Use quizzing with active retrieval of information at all phases of the learning process to exploit the ability of retrieval directly to facilitate long-lasting memory traces. 5a. Use pre-questions to introduce a new topic. 5b. Use quizzes to re-expose students to key content.</td>
<td>5a. Low 5b. Strong</td>
</tr>
<tr>
<td>6. Help students allocate study time efficiently. Assist students in identifying what material they know well, and what needs further study, by teaching children how to judge what they have learned. 6a. Teach students how to use delayed judgments of learning to identify content that needs further study. 6b. Use tests and quizzes to identify content that needs to be learned.</td>
<td>6a. Low 6b. Low</td>
</tr>
<tr>
<td>7. Ask deep explanatory questions. Use instructional prompts that encourage students to pose and answer “deep-level” questions on course material. These questions enable students to respond with explanations and supports deep understanding of taught material.</td>
<td>7. Strong</td>
</tr>
</tbody>
</table>
Checklist for carrying out the recommendations

**Recommendation 1:**
Space learning over time.
- Identify key concepts, terms, and skills to be taught and learned.
- Arrange for students to be exposed to each main element of material on at least two occasions, separated by a period of at least several weeks—and preferably several months.
- Arrange homework, quizzes, and exams in a way that promotes delayed reviewing of important course content.

**Recommendation 2:**
Interleave worked example solutions with problem-solving exercises.
- Have students alternate between reading already worked solutions and trying to solve problems on their own.
- As students develop greater expertise, reduce the number of worked examples provided and increase the number of problems that students solve independently.

**Recommendation 3:**
Combine graphics with verbal descriptions.
- Use graphical presentations (e.g., graphs, figures) that illustrate key processes and procedures. This integration leads to better learning than simply presenting text alone.
- When possible, present the verbal description in an audio format rather than as written text. Students can then use visual and auditory processing capacities of the brain separately rather than potentially overloading the visual processing capacity by viewing both the visualization and the written text.

**Recommendation 4:**
Connect and integrate abstract and concrete representations of concepts.
- Connect and integrate abstract and concrete representations of concepts, making sure to highlight the relevant features across all forms of the representation.

**Recommendation 5:**
Use quizzing to promote learning.
- Prepare pre-questions, and require students to answer the questions, before introducing a new topic.
- Use quizzes for retrieval practice and spaced exposure, thereby reducing forgetting.
- Use game-like quizzes as a fun way to provide additional exposure to material.

**Recommendation 6:**
Help students allocate study time efficiently.
- Conduct regular study sessions where students are taught how to judge whether or not they have learned key concepts in order to promote effective study habits.
- Teach students that the best time to figure out if they have learned something is not immediately after they have finished studying, but rather after a delay. Only after some time away from the material will they be able to determine if the key concepts are well learned or require further study.
- Remind students to complete judgments of learning without the answers in front of them.
- Teach students how to use these delayed judgments of learning techniques after completing assigned reading materials, as well as when they are studying for tests.
- Use quizzes to alert learners to which items are not well learned.
- Provide corrective feedback to students, or show students where to find the answers to questions, when they are not able to generate correct answers independently.

**Recommendation 7:**
Ask deep explanatory questions.
- Encourage students to “think aloud” in speaking or writing their explanations as they study; feedback is beneficial.
- Ask deep questions when teaching, and provide students with opportunities to answer deep questions, such as: What caused Y? How did X occur? What if? How does X compare to Y?
- Challenge students with problems that stimulate thought, encourage explanations, and support the consideration of deep questions.
Recommendation 1: **Space learning over time.**

To help students remember key facts, concepts, and knowledge, we recommend that teachers arrange for students to be exposed to key course concepts on at least two occasions—separated by a period of several weeks to several months. Research has shown that delayed re-exposure to course material often markedly increases the amount of information that students remember. The delayed re-exposure to the material can be promoted through homework assignments, in-class reviews, quizzes (see Recommendation 5), or other instructional exercises. In certain classes, important content is automatically reviewed as the learner progresses through the standard curriculum (e.g., students use single-digit addition nearly every day in second grade math), and this recommendation may be unnecessary in courses where this is the case. This recommendation applies to those (very common) course situations in which important knowledge and skills are not automatically reviewed.

**Level of evidence: Moderate**

The panel judges the level of evidence supporting this recommendation to be *moderate* based on three experimental classroom studies examining the effects of this practice for improving school-aged students’ performance on academic content (e.g., mathematics, spelling), 6 two experimental classroom studies that examined the effect of this strategy for improving college students’ academic performance, 7 and the hundreds of laboratory experiments which have been completed examining the effects of massed versus distributed practice on memory. 8

**Brief summary of evidence to support the recommendation**

Hundreds of laboratory experiments have been carried out which present materials to learners on two separate occasions. Then, following a delay, the learners are given some sort of recall test on the material. Although a few inconsistencies have been found, by far the most common finding is that when the time between study sessions is very brief relative to the amount of time to the final test, students do not do as well on the final test. 9 Students typically remember much more when they have been exposed to information on two occasions, rather than one, and when the interval between these two occasions is not less than about 5 percent of the interval during which the information has to be retained. In the studies that have tested this principle of delayed review, researchers have kept constant the amount of time that students have to learn the information; 10 thus, the observed improvement in learning is not a result of learners having more time to study the material. Delaying of reviews produces an actual increase in the efficiency of learning. Having too long a temporal spacing separating learning sessions has been found to produce a small decrease in final memory performance as compared to an optimal spacing, but the cost of “overshooting” the right spacing is consistently found to be much smaller than the cost of having very short spacing. Thus, the practical implication is that it makes sense to be sure to have enough spacing, but it rarely makes sense to worry about having too much.

---

7 Rohrer and Taylor (2006); Bahrick, Bahrick, Bahrick, et al. (1993).
9 Examples of what is meant by a brief interval relative between study sessions would be a 10-second interval when the test occurs a half hour later, or a one-day delay when the test occurs months later.
10 For example, one group of students might spend 20 minutes learning the definitions of a list of words and then have a test on those words ten days later. These students would be compared to a group of students who spend 10 minutes on one day learning the definitions and then 10 minutes on another day reviewing the definitions.
Research on the delayed review of materials has examined learning of (a) mathematical skills,\textsuperscript{11} (b) foreign language vocabulary,\textsuperscript{12} and (c) historical and other facts.\textsuperscript{13} Although the research literature primarily involves well-controlled laboratory studies, there are a number of classroom-based studies that have shown similar results. One recent study examined memory for historical facts by eighth-graders enrolled in a U.S. history class.\textsuperscript{14} The study compared the effect of a review given 1 week after initial presentation, versus 16 weeks after. On a final test given 9 months after the review session, the 16-week delay group showed significantly greater performance (almost 100 percent increase) as compared to the 1-week delayed group.

One limitation of the literature is that few studies have examined acquisition of complex bodies of structured information.\textsuperscript{15} For measurement reasons, researchers have mostly focused on acquisition of isolated bits of information (e.g., facts or definitions of vocabulary words). The acquisition of facts and definitions of terms is certainly an essential component of mastering any complex content domain, and may have broad cultural utility,\textsuperscript{16} but the panel recognizes that acquiring facts and key definitions is merely one goal of schooling.\textsuperscript{17} There does not appear to be any evidence to suggest that spacing benefits are confined to isolated elements of course content.

How to carry out the recommendation

The key action recommended here is for teachers to make sure that important curriculum content is reviewed at least several weeks, and ideally several months, after the time that it was first encountered by the students. Research shows that a delayed review typically has a large positive impact on the amount of information that is remembered much later. The benefit of a delayed review seems to be much greater than the same amount of time spent reviewing shortly after initial learning. This review can occur in a variety of ways, including those described below.

1. **Use class time to review important curriculum content.**

For example, every other week a high school social studies teacher spends half a class period reviewing facts that were covered several weeks earlier in the class.

2. **Use homework assignments as opportunities for students to have spaced practice of key skills and content.**

For example, in every homework assignment, a junior high school math teacher intentionally includes a few problems covering the material presented in class 1 or 2 months earlier.

3. **Give cumulative midterm and final examinations.**

When teachers give their students cumulative midterm and final examinations, students are provided with a strong incentive to study all course material at widely separated points in time.

Possible roadblocks and solutions

**Roadblock 1.1.** Most textbooks contain reviews and problem sets that deal only with the most recently taught material.

**Solution.** Teachers can supplement problem sets provided in the textbook with at least a “sprinkling” of problems relating to material covered much earlier in the course. One may hope that in the future, textbook publishers will respond to the growing body of research on spacing of learning and develop textbooks that directly promote spaced review of key concepts and procedures.

**Roadblock 1.2.** Teachers may frequently become discouraged during a review session to discover that many students appear to have forgotten what they appeared to have mastered several weeks earlier.

**Solution.** By implementing our recommended practice of spacing over time, teachers will find that students

\textsuperscript{11} E.g., Rohrer and Taylor (2006, in press).
\textsuperscript{12} E.g., Dempster (1987); Bahrick, Bahrick, Bahrick, et al. (1993).
\textsuperscript{13} E.g., Carpenter, Pashler, Cepeda, et al. (2007); Pashler, Rohrer, Cepeda, et al. (2007).
\textsuperscript{14} Carpenter, Pashler, Cepeda, et al. (2007).
\textsuperscript{15} Ausubel and Youssef (1965) showed benefits of delayed review on memory for a coherent passage on endocrinology, but the comparison was with a procedure that lacked the delayed review (rather than one that included a review at a short lag).
\textsuperscript{16} See Hirsch (1987) for a discussion.
\textsuperscript{17} See Bloom (1956) for a well-known discussion.
remember more. At the beginning of this process, teachers should expect to see substantial forgetting. Although this initial forgetting may be discouraging, the panel reminds our readers that research shows that even when students cannot recall previously learned material, reawakening of the knowledge through reviewing is more easily accomplished than was the original learning (psychologists refer to this as “savings”), and the final result of the delayed review is a marked reduction in the rate of subsequent forgetting.\(^\text{18}\) Thus, by implementing spaced review, the teacher can not only repair the forgetting that will have happened since initial learning, but also, to some degree, inoculate against subsequent forgetting.

\(^{18}\) Berger, Hall, and Bahrick (1999).